



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 119902

TO: **Phuong N Huynh**
Location: 3d75, 3c70
Saturday, April 24, 2004
Art Unit: 1644
Phone: 272-0846
Serial Number: 10 / 005907

From: Jan Delaval
Location: Biotech-Chem Library
Rem 1A51
Phone: 272-2504
jan.delaval@uspto.gov

Search Notes

Best Local Similarity: 85.19% Mismatches: 0
 Query Match: 81.68% Indels: 20
 DB: 9 Gaps: 1

US-10-005-907-2 (1-135) x AK126682 (1-4032)

QY 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGluGlnGlnLysProLys 20
 DB 377 ATGGGAATATCTCTCGGAATCTC-----403

QY 21 LysGlyAsnProAspGluGluArgLysGlnGluMetThrThrPheGluArgLysLeu 40
 DB 404 -----AGGCAGGAATGACTACATTTGAAAGAAACTT 436

QY 41 GlnAspGlnAspLysSerGlnGluValSerThrSerAsnGlnGluGlnGluAsn 60
 DB 437 CAGAGAGATATGAAAGGCGAGATTTCCTCTTATTCAGGAAACGAGAT 496

QY 61 GlySerGlySerGluGluValCysTyrThrValIleAsnHisIleProHisGlnArgSer 80
 DB 497 GGCAGTCTATCTGAGGAAGTGTGTACACTGTCTATTATCATCATCCCATCAGAGATCC 556

QY 81 SerLeuSerSerAsnAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArg 100
 DB 557 TCCCTGAGCTTCAATGATGATGGCTATGAGAACATTCCTCCCTCACAAGGAAGTGAGA 616

QY 101 GlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArgPro 120
 DB 617 CAGTTTACAGAAAGTCTCAGACAGACAGATATGCCCTTCTTAGGACTTCTGTAGTAGGCT 676

QY 121 CysSerCysThrHisGluHisAspTyrGluValValPheProHis 135
 DB 677 TGTTCCTGCACCCATGACATGATTATGAGTTGTGTGTTCACAC 721

RESULT 6
 AK124520 4136 bp mRNA linear
 LOCUS Homo sapiens cDNA FLJ43529 fis, clone BRACE3002390.
 DEFINITION AK124520
 VERSION AK124520.1 GI:34530324
 KEYWORDS oligo capping; fis (full insert sequence).
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
 AUTHORS Ninomiya, K., Wagatsuma, M., Kanda, K., Kondo, H., Yokoi, T., Kodaira, H., Furuya, T., Takahashi, M., Kikkawa, E., Omura, Y., Abe, K., Kamihara, K., Katuta, N., Sato, K., Tanikawa, M., Yamazaki, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamae, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Murakawa, K., Kanekuri, K., Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahara, K., Masuho, Y., Nagai, K. and Itoga, T.
 NEDO human cDNA sequencing project
 Unpublished
 2 (bases 1 to 4136)
 Isogai, T. and Yamamoto, J.
 Direct Submission
 Submitted (15-JUN-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7
 Karusa-Kamatari, Kisarazu, Chiba 292-0818, Japan
 (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
 NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology (RAB); CDNA library construction; Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing; RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing; HRI and RAB; annotation; HRI and RAB.
 Location/Qualifiers
 1..4136

US-10-005-907-2 (1-135) x AK124520 (1-4136)

QY 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGluGlnGlnLysProLys 20
 DB 498 ATGGGAATATCTCTCGGAATCTC-----524

QY 21 LysGlyAsnProAspGluGluArgLysGlnGluMetThrThrPheGluArgLysLeu 40
 DB 525 -----AGGCAGGAATGACTACATTTGAAAGAAACTT 557

QY 41 GlnAspGlnAspLysSerGlnGluValSerThrSerAsnGlnGluGlnGluAsn 60
 DB 558 CAGATATCAGATATGAGAAAGCGAGAGATTTCATCTTATTCAGGAAACGAGAT 617

QY 61 GlySerGlySerGluGluValCysTyrThrValIleAsnHisIleProHisGlnArgSer 80
 DB 618 GGCAGTGTGTCTGAAAGAGTGTGTACACTGTCTATTATTCATCCCATCAGAAATCC 677

QY 81 SerLeuSerSerAsnAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArg 100
 DB 678 TCCCTGAGCTTCAATGATGATGGCTATGAGAACATTCCTCCCTCACAAGGAAGTGAGA 737

QY 101 GlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArgPro 120
 DB 738 CAGTTTACAGAAAGTCTCAGACAGACAGATATGCCCTTCTTAGGACTTCTGTAGTAGGCT 797

QY 121 CysSerCysThrHisGluHisAspTyrGluValValPheProHis 135
 DB 798 TGTTCCTGCACCCATGACATGATTATGAGTTGTGTGTTCACAC 842

RESULT 7
 AC074365/c 141268 bp DNA linear
 LOCUS Homo sapiens chromosome 1 clone RP11-115C4, WORKING DRAFT SEQUENCE.
 DEFINITION AC074365
 VERSION AC074365.5 GI:10280935
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 141268)
 AUTHORS Waterston, R.H.
 TITLE The sequence of Homo sapiens clone
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 141268)
 AUTHORS Waterston, R.H.

http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr1>
 RP11-978115 is from the library RPC1-11.4 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>
 VECTOR: pBACe3.6

----- Genome Center
 Center: Wellcome Trust Sanger Institute
 Center code: SC
 Web site: <http://www.sanger.ac.uk>
 Contact: humquery@sanger.ac.uk

FEATURES

Location/Qualifiers

1..185467
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="1"
 /clone="RP11-978115"
 /clone_lib="RPC1-11.4"

ORIGIN

Alignment Scores:
 Pred. No.: 2,13e-33 Length: 185467
 Score: 430.00 Matches: 80
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 60.14% Indels: 0
 DB: 9 Gaps: 0

US-10-005-907-2 (1-135) x AL606804 (1-185467)

QY 56 GlnGluAsnGluAsnGlySerGlySerGluGluValCysThrValIleAsnHisIle 75
 DB 132080 CAGGAAACGAGATGCGAGTGGTCTGAAGAGTGTGCTACATGCTCAATATCACATC 132139
 QY 76 ProHisGlnArgSerSerLeuSerSerAsnAspAspGlyThrGluAsnIleAspSerLeu 95
 DB 132140 CCCATCAGAGATCCCTCCCTGAGCTCCATGATGATGGGTATGAGAACATGACCTCCCTC 132199
 QY 96 ThrArgLysValArgGlnPheArgGluArgSerGluThrGluThrGluThrValAlaLeuLeuArgThr 115
 DB 132200 ACAGGAAGATGACAGCAGTTAGAGAAAGTCGACAGCAGATATGCCCCCTCTTAGGACT 132259
 QY 116 SerValSerArgProCysSerCysThrHisGluHisAspTyrGluValValPheProHis 135
 DB 132260 TCTGTTAGTAGGCGCTTCTCTCCAGCCATGAGCATGATATGAGTTGTTGTTCCACAC 132319

RESULT 9
 BC030506
 LOCUS
 DEFINITION Homo sapiens germinal center expressed transcript 2, mRNA (CDNA clone MGC:40441 IMAGE:4385178), complete cds.

ACCESSION BC030506.1 GI:20987805
 VERSION BC030506
 KEYWORDS MGC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 1532)
 AUTHORS Strausberg, R.L., Feingold, R.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altshuler, S.F., Zeeberg, K.H., Sutton, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Mariani, K., Parker, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Uebachs, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J.,

McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, D.J., Hulyk, S.W., Villalón, D.C., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dikson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Smallos, D.E., Schnarch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 TITLE
 JOURNAL MEDLINE
 PUBLISHED
 REFERENCE
 2. (bases 1 to 1532)
 Strausberg, R.
 Direct Submission
 Submitted (07-MAY-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabbs-remail@nih.gov
 Tissue Procurement: You Staudt
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
 DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC)
 Gaithersburg, Maryland;
 Web site: <http://www.nisc.nih.gov/>
 Contact: nisc.mgc@nih.gov
 Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Gupta, J., Haghighi, P., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, O.L., Masiello, C., Maskari, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stanke, P., Thomas, P.J., Touchman, J.W., Teague, A., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 64 Row: a Column: 15
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 22749536.

FEATURES

Location/Qualifiers

1..1532
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:40441 IMAGE:4385178"
 /tissue_type="Lymph, lymphoma"
 /clone_lib="NIH MGC 85"
 /lab_host="DH10B"
 /note="vector: pCMV-SPORT6"
 1..1532
 /gene="GCET2"
 /note="synonyms: HGAL, GCAT2, MGC40441"
 /db_xref="LocusID:257144"
 /db_xref="MIN:607792"
 54..590
 /codon_start=1
 /product="germinal center expressed transcript 2"
 /protein_id="AAH30506.1"
 /db_xref="GI:20987806"
 /translation="MGSLRLNRKQNTQMPVNVNQSPKORTSCWDHHAAGCCPCLPWKILIFSKQDSQNERMSSTPIQNVDTQSEELCYTLINRVLCTPSPQMSASBYENVPKAEPRSLSGTETISLLHMPSTDPRHASPEDVELLAPHRISSHP LQQRPLMAPSETQFSLH"

gene

CDS

```

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE Bull, J.H., Ellison, G. and Paskins, L.D.
AUTHORS Diagnostic methods for the detection of prostatic disorders
TITLE Patent: WO 0136674-A 20-MAY-2001;
JOURNAL AstraZeneca AB (SE)
FEATURES Location/Qualifiers
source 1..1933
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"
ORIGIN
Query Match 1.3%; Score 49; DB 6; Length 1933;
Best Local Similarity 100.0%; Fred.No. 5.6e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3714 AATCTCAAAATGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 3762
DB 1878 AATCTCAAAATGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1926
RESULT 36
HUMSAP1A Homo sapiens 1933 bp mRNA linear PRI 17-DEC-1993
LOCUS HUMSAP1A Homo sapiens SRP accessory protein 1A (SAP-1) mRNA, complete cds.
DEFINITION HUMSAP1A
ACCESSION M85165
VERSION M85165.1 GI:429185
KEYWORDS serum response factor; SAP-1; Elk-1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1933)
Dalton.S. and Treisman,R.
Characterization of SAP-1, a protein recruited by serum response
factor to the c-fos serum response element
Cell 68 (3), 597-612 (1992)
JOURNAL MEDLINE 92154673
PUBMED 1339307
REFERENCE 2 (bases 1 to 1933)
Treisman,R.
Direct Submission
TITLE Submitted (05-MAR-1992) Richard Treisman, Transcription Laboratory,
Imperial Cancer Research Fund, London, England
COMMENT On Nov 29, 1993 this sequence version replaced gi:338036.
Original source text: Homo sapiens CDNA to mRNA.
Location/Qualifiers
1..1933
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
1..1933
    /gene="SAP-1"
    /note="Homology region A with Elk-1 protein (Ets domain)
is bp 150-417, amino acids 1-89; Homology region B with
Elk-1 protein, required for cooperative ternary complex
formation with SRF is bp 565-617, amino acids 136-157;
Sequence diverges from SAP-1B at bp 1229, amino acid 360;
Homology region C with Elk-1 protein, core of regulated
transcription activation domain, is bp 1203-1355, amino
acids 352-402; (SVT)P motifs conserved between SAP-1A and
Elk-1 are located at amino acids T354, T361-T366, S381,
S387, T420, S425, corresponding to bp 1209-1214, bp
1230-1235, bp 1243-1250, bp 1290-1295, bp 1308-1313, bp
1407-1412, bp-1422-1427"
/codon_start=1
/product="SAP-1A protein"
/protein_id="AAA03631.1"
FEATURES source
gene
CDS

```

FEATURES
Source Location/Qualifiers
1. .527
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4109064"
/tissue_type="from acute myelogenous leukemia"
/lab_hosts="DH10B (TI phage-resistant)"
/clone_lib="NIH_MGC_55"
/note="Organ: bone marrow; Vector: pDNR-LIB (Clontech); Site 1: SfiI (ggcgcttcggcc); Site 2: SfiI (ggcattatggcc); Double-stranded cDNA was prepared from cell line RNA. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGCCCATATGACC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCGGCGGCGGCGCATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.65 kb (range 0.9-4.0 kb). 14/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

ORIGIN
Query Match 100.0%; Score 405; DB 10; Length 527;
Best Local Similarity 100.0%; Pred. No. 6.1e-86;
Matches 405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGAAATATCTCTCGGAAACTCAGTTGCTGGGAGAGATCAAAAGAGCCCAAG 60
DB 45 ATGGGAAATATCTCTCGGAAACTCAGTTGCTGGGAGAGATCAAAAGAGCCCAAG 104
QY 61 AAGGAAACCCAGATGAGAAAGAAACCGGAGGAAATGACTACATTTGAAAGAAACTT 120
DB 105 AAGGAAACCCAGATGAGAAAGAAACCGGAGGAAATGACTACATTTGAAAGAAACTT 164
QY 121 CAGCATCAAGATAGAAAGCCAGAGCTTTCATCCACTTCTATATCAGGAAACGAGAT 180
DB 165 CAGCATCAAGATAGAAAGCCAGAGCTTTCATCCACTTCTATATCAGGAAACGAGAT 224
QY 181 GGCAGTGGTTCTGAAGAGTGTGTACACTGTCTATTAATCATCCCTCCCATCAGATCC 240
DB 225 GGCAGTGGTTCTGAAGAGTGTGTACACTGTCTATTAATCATCCCTCCCATCAGATCC 284
QY 241 TCCCTGAGCTCCATGATGATGGTGTAGACATGATGATCCCTCCATCAGGAAAGTGAGA 300
DB 285 TCCCTGAGCTCCATGATGATGGTGTAGACATGATGATCCCTCCATCAGGAAAGTGAGA 344
QY 301 CAGTTAGAGAAAGCTCAGAGACAGAAATATGCCCTTCTTAGGACTTCTGTAGAGCCT 360
DB 345 CAGTTAGAGAAAGCTCAGAGACAGAAATATGCCCTTCTTAGGACTTCTGTAGAGCCT 404
QY 361 TGTTCCTGCCACCCATGAGCATGATGAGAGTGTGTGTTCACAC 405
DB 405 TGTTCCTGCCACCCATGAGCATGATGAGAGTGTGTGTTCACAC 449

RESULT 2
AW480906
LOCUS AW480906 288 bp mRNA linear EST 09-JUL-2000
DEFINITION 3516 MARC 2P1G Sus scrofa cDNA 5', mRNA sequence
ACCESSION AM480906
VERSION AM480906.1 GI:7050949
KEYWORDS EST.
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
1 (Bases 1 to 288)
Fahrenkrug, S.C., Smith, T.P.L., Freking, B.A., Cho, J., White, J.,
Vallet, J., Wiss, A., Rohrer, G.A., Pertea, G., Sultana, R.,
Quackenbush, and Keele, J.W.
Porcine genome discovery by normalized cDNA-library sequencing and
EST cluster assembly

Sat Apr 24 08:29:31 2004

us-10-005-907-1.oligo.rge

```
/db_xref="GI:429186"
/translation="MDSAITLWQFLQLLQKPKNKHMCWTSNDGQFKLQAEVRL
WGIKKNPNMYDKLSALRYVYVQNIKKVNGOKFVYKVPYPIILNMDPMTVGRIE
GDCESLNPSVSSXDVENGCKPQPGAKTSRNDYIHSGLYSSFTLNSLNSNV
KLFKLIETNPRAEKLAEKKQPOPTSVIKVTTSPKPKVPEVPAATISIGSISPS
BETIOALETVSPKLELEAPTSASNMVPAATTPRISSIDPLOEPPRTSPPLSHHP
DIDTIDSVASQPMELFENLSLEPKQDSVLEKKNVNSRSKPKGKGLATLVIT
SSDEPLGLSPSLPTASLTPAFFSQTPILUTPSELLSSIHFWSTLSPVAPLSAPLQ
CANTLFOFPVLSNHRGPTTSLGMDGPESTEGFPFDLQKT"
```

ORIGIN

```
Query Match 1.3% Score 49; DB 9; Length 1933;
Best Local Similarity 100.0%; Pred. No. 5.6e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 3714 AATGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
|||
Db 1878 AATGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1926
```

BLASTN-67

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 22, 2004, 13:36:26 ; Search time 2525 Seconds
(without alignments)
4789.775 Million cell updates/sec

Title: US-10-005-907-1_COPY_25_429

Perfect score: 405
Sequence: 1 atgggaattattctctgcg.....atgaagtgtgtttccacac 405

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_estc:*

9: gb_est1:*

10: gb_est2:*

11: gb_est3:*

12: gb_est4:*

13: gb_est5:*

14: gb_est6:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_phg:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	405	100.0	527	10	BF242113
2	64.8	16.0	288	10	AM480905
3	59.6	14.7	1045	10	BF797507
4	58.6	14.5	1085	12	EM455407

5	58	14.3	508	29	AY404586
6	58	14.3	876	13	BQ214406
7	58	14.3	924	12	BG389415
8	58	14.3	987	9	AL560376
9	58	14.3	1000	12	BM904106
10	58	14.3	1039	12	BM479887
11	58	14.3	1070	12	BM456595
12	58	14.3	1095	12	BG106563
13	58	14.3	1124	10	BF795688
14	58	14.3	1193	12	BM455198
15	57.4	14.2	526	14	CB286394
16	56.4	13.9	683	10	AM963366
17	56.2	13.9	1165	12	BM458339
18	48.4	12.0	495	9	AA311038
19	47.2	11.7	1041	10	BG024891
20	47	11.6	1059	12	BM459235
21	46.8	11.6	353	29	AY404587
22	46.6	11.5	259	29	CE192340
23	45.4	11.2	658	29	CE020416
24	45	11.1	967	14	CD105480
25	44.8	11.1	826	13	BU746391
26	44.8	11.1	840	13	BU746392
27	44.6	11.0	560	13	BU695543
28	44.6	11.0	774	28	SI9721
29	44.4	11.0	731	29	CE188479
30	44	10.9	203	29	CE091328
31	44	10.9	421	28	AQ501175
32	44	10.9	520	12	BG145241
33	44	10.9	522	12	BG149093
34	44	10.9	524	12	BG276802
35	44	10.9	526	12	BG278546
36	44	10.9	668	29	CE207791
37	44	10.9	825	14	CB320849
38	43.8	10.8	899	29	CNS03004
39	43.8	10.8	959	29	CNS00655
40	43.8	10.8	1201	9	AL514529
41	43.6	10.8	815	13	BU932945
42	43.4	10.7	404	13	EX367419
43	42.8	10.6	703	29	CE790536
44	42.8	10.6	756	29	CE196221
45	42.6	10.5	891	10	BF264660

ALIGNMENTS

RESULT 1
BF242113
LOCUS 601880401F1_MTH_MGC_55 Homo sapiens cDNA clone IMAGE:4109064 5',
DEFINITION 527 bp mRNA linear EST 14-NOV-2000
ACCESSION BF242113
VERSION BF242113.1 GI:11156040
KEYWORDS EST.
SOURCE Homo sapiens
ORGANISM Homo sapiens (human)
REFERENCE 1 (bases 1 to 527)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: CLONETECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LCM993 row: a column: 01
High quality sequence stop: 518.


```
FEATURES
  source
    Location/Qualifiers
      1..527
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="IMAGE:4109064"
        /tissue_type="from acute myelogenous leukemia"
        /lab_host="DH10B (T1 phage-resistant)"
        /clone_lib="NIH MGC 55"
        /note="Organ: bone marrow; Vector: pDNR-LIB (Clontech);
        Site 1: SfiI (ggcgctcgcc); Site 2: SfiI
        (ggcattatggcc); Double-stranded cDNA was prepared from
        cell line RNA. 5' and 3' adaptors were used in cloning as
        follows: 5' adaptor sequence: 5'-CACGGCCATTATGCC-3' and
        3' adaptor sequence:
        5'-ATTCAGAGCGCGGCGGCGGCACATG-dT(30)EN-3' (where B = A,
        C, or G and N = A, C, G, or T). Average insert size
        1.65 kb (range 0.9-4.0 kb). 14/15 colonies contained
        inserts by PCR. This library was enriched for full-length
        clones and was constructed by Clontech Laboratories (Palo
        Alto, CA)."
      ORIGIN
        Query Match 100.0%; Score 405; DB 10; Length 527;
        Best Local Similarity 100.0%; Pred. No. 6.1e-86;
        Matches 405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
        QY 1 ATGGGAATATTCTCTCTGCGAAATCTCAGTTGCTGCGAGAGAAATCAAAAGAGCCCAAG 60
        DB 45 ATGGGAATATTCTCTCTGCGAAATCTCAGTTGCTGCGAGAGAAATCAAAAGAGCCCAAG 104
        QY 61 AAAGGAACCCAGATGAGGAGAGAAAGCGGAGGAAATGACTTATGAAAGAAATCTT 120
        DB 105 AAAGGAACCCAGATGAGGAGAGAAAGCGGAGGAAATGACTTATGAAAGAAATCTT 164
        QY 121 CAAGATCAAGATAGAGAAAGCCAGAGGTTTCATCCACTTCTTAATCAGGAAACGAGAAAT 180
        DB 165 CAAGATCAAGATAGAGAAAGCCAGAGGTTTCATCCACTTCTTAATCAGGAAACGAGAAAT 224
        QY 181 GCAGTGGTTCGAGAGAGTGTCTACATGCTATTAATCATCCATCCATCCATCCATCCATCC 240
        DB 225 GCAGTGGTTCGAGAGAGTGTCTACATGCTATTAATCATCCATCCATCCATCCATCCATCC 284
        QY 241 TCCTGAGCTCCAATGATGCTATGAGACATGATGCTATGAGACATGCTATGAGACATGCTATGAG 300
        DB 285 TCCTGAGCTCCAATGATGCTATGAGACATGATGCTATGAGACATGCTATGAGACATGCTATGAG 344
        QY 301 CAGTTAGAGAAAGGTCAGAGACAGAAATATGCCCTCTTAGGACTTCTGTAGTAGGCTT 360
        DB 345 CAGTTAGAGAAAGGTCAGAGACAGAAATATGCCCTCTTAGGACTTCTGTAGTAGGCTT 404
        QY 361 TGTCTCTGCCCATGAGCATGATTAAGTGTGTTTCCACAC 405
        DB 405 TGTCTCTGCCCATGAGCATGATTAAGTGTGTTTCCACAC 449
      RESULT 2
      AW480906
      LOCUS 33516 MARC 2P1G Sus scrofa cDNA 5', mRNA linear EST 09-JUL-2000
      DEFINITION
      ACCESSION AW480906
      VERSION AW480906.1 GI:7050949
      KEYWORDS EST.
      SOURCE Sus scrofa (pig)
      ORGANISM Sus scrofa
        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
        Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
      REFERENCE 1 (bases 1 to 288)
      AUTHORS Fahrtenkrug,S.C., Smith,T.P.L., Fraking,B.A., Cho,J., White,J.,
        Vallet,J., Wise,T., Rohrer,G.A., Perle,G., Sultana,R.,
        Quackenbush,J. and Keele,J.W.
      TITLE Porcine gene discovery by normalized cDNA-library sequencing and
        EST cluster assembly
      Mamm. Genome 13 (8), 475-478 (2002)
      JOURNAL 22213789
      MEDLINE 12226715
      PUBMED
      COMMENT
        Contact: Smith TPL
        USDA, ARS, US Meat Animal Research Center
        PO Box 166, Clay Center, NE 68933-0166, USA
        Tel: 402 762 4366
        Fax: 402 762 4390
        Email: smith@mail.marc.usda.gov
        Single pass sequencing. Bases called and trimmed with phred
        v0.980904.e. Vector identified by cross_match with the -minscore 20
        and -mismatch 12 options.
        PCR Primers
        FORWARD: AGGAAACAGCTATGACCAT
        BACKWARD: GTTTCCTCAGTCACGACG
        Plate: 19 row: F column: 21
        Seq primer: ATTAGGTGACACTATAG.
      FEATURES
        source
          1..288
            /organism="Sus scrofa"
            /mol_type="mRNA"
            /db_xref="taxon:9823"
            /tissue_type="pooled"
            /lab_host="DH10B"
            /clone_lib="MARC 2P1G"
            /note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
            Library made from pooled tissue from testis, ovary,
            endometrium, hypothalamus, pituitary, and placenta."
          ORIGIN
            Query Match 16.0%; Score 64.8; DB 10; Length 288;
            Best Local Similarity 75.8%; Pred. No. 4e-05;
            Matches 94; Conservative 0; Mismatches 27; Indels 3; Gaps 1;
            QY 285 CACAAGAAAGTCAGACAGCTTAGAGAAAGTCAGAGACAGAAATATGCCCTTCTTAGGAC 344
            DB 2 CACAAGAGAGTCAGACCACTTAAGAGGAGATCAGAAACAGAAATATGCCCTTCTTAGGAC 61
            QY 345 TTCTGTAGT---AGGCTTGTTCCTGACCATGAGCATGATTATGAAGTGTGTTTCC 401
            DB 62 GATTACATCCAGTTCCTTCTCTATACCTGAGATGATTATGAAGTGTGTTTCC 121
            QY 402 ACAC 405
            DB 122 TCAC 125
          RESULT 3
          BF797507
          LOCUS 602257174F1 NIH MGC_85 Homo sapiens cDNA clone IMAGE:4340323 5',
            mRNA sequence.
          DEFINITION
          ACCESSION BF797507
          VERSION BF797507.1 GI:12102561
          KEYWORDS EST.
          SOURCE Homo sapiens (human)
          ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
          REFERENCE 1 (bases 1 to 1045)
          AUTHORS NIH-MGC http://mgi.nci.nih.gov/
          TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
          JOURNAL Unpublished (1999)
          COMMENT
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs@mail.nih.gov
            Tissue Procurement: Louis Staudt, M.D., Ph.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: L1AM952 row: d column: 20
```

FEATURES
Source
High quality sequence stop: 687.
Location/Qualifiers
1..1045

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4340323"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_85"
/note="Organ: lymph; Vector: pCMV-SPORT6; Site: 1: NotI; Site: 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.867 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 14.7%; Score 59.6; DB 10; Length 1045;
Best Local Similarity 54.7%; Pred. No. 0.00075;
Matches 163; Conservative 0; Mismatches 129; Indels 6; Gaps 2;

QY 114 AAAAATTCAGATCAAGATGAAGAAAGCCCAAGAGTTTCATCCACTTCTAATCAGGAAAA 173
DB 267 AAAGAGGCAAGATTCCTCCAAACCAAAATGAAGAAATGTCATCTACTCTCCATCCAGGACAA 326

QY 174 CGAGATGGCAGTGGTCTGAGAGAGTGCTGCTACCTGTCATTAATCACA---TCCCCCA 230
DB 327 TGTGACCAAGCTACTCAGAGAGCTGTGCTATACCTCATCATATCGGGTCTCTG 386

QY 231 TCAGAGATCTCTCCCTGAGCTCCAAATGATGAGTATGAGAACTGATGCTCCCTCACAAG 290
DB 387 TACAAGGCCATCAGGGAACCTCTGCTGAGAGAGTACTATGAGATGTTTCCCTGCAAGCTGA 446

QY 291 GAAAGTGAGACAGTTAGAGAAAGCTCAGAGACAGATATGCCCTTCT---TAGGACTTC 347
DB 447 GAGACCCAGAGTCTCTGAGAGAACTGAGCTGAGTATCTACTTACATATGCTTC 506

QY 348 TGTAGTAGGCTTGTCTGACCCATGAGCATGATGAGATGAGTGTGTTTCCACAC 405
DB 507 TACAGACCCAGGATGCCGATCCCGAGAGAGATGATATGAACTTCTCATGCTCAC 564

RESULT 4
BM455407
LOCUS BM455407.1 1085 bp mRNA linear EST 05-FEB-2002
DEFINITION AGENCOURT_6407334 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:5500391
5', mRNA sequence.
ACCESSION BM455407
VERSION BM455407.1 GI:18504436
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Straubeberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Lou Staudt
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM12135 row: d column: 24
High quality sequence start: 19
High quality sequence stop: 760.
Location/Qualifiers
1..1085
/organism="Homo sapiens"

FEATURES
source

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5500391"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_85"
/note="Organ: lymph; Vector: pCMV-SPORT6; Site: 1: NotI; Site: 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.867 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 14.5%; Score 58.6; DB 12; Length 1085;
Best Local Similarity 54.4%; Pred. No. 0.0013;
Matches 162; Conservative 0; Mismatches 130; Indels 6; Gaps 2;

QY 114 AAAAATTCAGATCAAGATGAAGAAAGCCCAAGAGTTTCATCCACTTCTAATCAGGAAAA 173
DB 459 AAAGAGCAAGATTCCTCCAAACCAAAATGAAGAAATGTCATCTACTCTCCATCCAGACAA 518

QY 174 CGAGATGGCAGTGGTCTGAGAGAGTGCTGCTACCTGTCATTAATCACA---TCCCCCA 230
DB 519 TGTGACCAAGCTACTCAGAGAGCTGTGCTATACCTCATCATATCGGGTCTCTG 578

QY 231 TCAGAGATCTCTCCCTGAGCTCCAAATGATGAGTATGAGAACTGATGCTCCCTCACAAG 290
DB 579 TACAGGCCATCANGAACTCTCTGAGAGTACTATGAGATGTTTCCCTGCAAGCTGA 638

QY 291 GAAAGTGAGACAGTTTAGAGAAAGCTCAGAGACAGATATGCCCTTCT---TAGGACTTC 347
DB 639 GAGACCCAGAGATCTCTGAGAGAACTGAGCTGAGTATCTACTTACATATGCTTC 698

QY 348 TGTAGTAGGCTTGTCTGACCCATGAGCATGATGAGTGTGTTTCCACAC 405
DB 699 TACAGACCCAGGATGCCGATCCCGAGAGATGATATGAACTTCTCATGCTCAC 756

RESULT 5
AY404586
LOCUS AY404586 508 bp DNA linear GSS 16-DEC-2003
DEFINITION Homo sapiens GCET2 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.
ACCESSION AY404586
VERSION AY404586.1 GI:39760563
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A., Todd, M.A., Tanenbaum, D.M., Civallo, D.R., Lu, F., Murphy, B., Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J., Adams, M.D. and Cargill, M.
Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
Science 302 (5652), 1960-1963 (2003)
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 508)
AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A., Todd, M.A., Tanenbaum, D.M., Civallo, D.R., Lu, F., Murphy, B., Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J., Adams, M.D. and Cargill, M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering them based on alignment.
Location/Qualifiers
1..508
/organism="Homo sapiens"
/mol_type="genomic DNA"

```

/db xref="taxon:9606"
<1..2508
/seq="GCEr2"
/locus_tag="HCM1940"

ORIGIN
Query Match      14.3%; Score 58; DB 29; Length 508;
Best Local Similarity 54.4%; Pred. No. 0.0017;
Matches 162; Conservative 0; Mismatches 130; Indels 6; Gaps 2;

QY 114 AAAAATTCAGATCAAGATGAAGAAAGCCAGAGCTTCATCCACTTCAATCAGGAAA 173
DB 136 AAAGAGCAGATTCCTCAAAACGAAATGAAGATGCTATCTCTCCATCCAGGAAA 195
QY 174 CGAAGTGGCAGTGGTTCGAGAGAGTGTGTACACTGTCACTTCAATATCA---TCCCCCA 230
DB 196 TGTGACACAGACCTACTCAGAGAGCTGTGTCTATACCTCTCAATCATCGGGTCTCTG 255
QY 231 TCAGAGATCTCTCCTCAGCTTCAATGATGCTGATGAGAACTTGAATGCTCTCCATCA 290
DB 256 TACAGGCCATCAGGAACTCTGCTGAGAGTACTATGAGATGTTCCCTGCAAGCTGA 315
QY 291 GAAAGTGAACAGTTAGAGAAAGGTCAGAGACAGAAATATGCCCTTCT---TAGGACTTC 347
DB 316 GAGACCCAGAGATCTCTGAGAGAACTGAGACTGAGTATTCATCTTACATATATGCCCTTC 375
QY 348 TGTAGTAGGCTTGTCTCTGACCCATGAGCATGATTAATGAGTGTGTGTTCCACAC 405
DB 376 TACAGACCCAGGAGTCCCGATCCCGAGAGATGAATGATGAATCTCTCATGCTCTCAC 433

RESULT 6
BQ214406
LOCUS AGENCOURT_7592611 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:6070069
5', mRNA sequence.
DEFINITION BQ214406.1 GI:20395806
ACCESSION BQ214406
VERSION BQ214406
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 876)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13353 row: m column: 14
High quality sequence stop: 678.
FEATURES
Location/Qualifiers
1..876
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6070069"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_92"
/notes="Organ: testis; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI. Cloned unidirectionally; oligo-dT primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
ORIGIN

```

```

Query Match      14.3%; Score 58; DB 13; Length 876;
Best Local Similarity 54.4%; Pred. No. 0.0018;
Matches 162; Conservative 0; Mismatches 130; Indels 6; Gaps 2;

QY 114 AAAAATTCAGATCAAGATGAAGAAAGCCAGAGCTTCATCCACTTCAATCAGGAAA 173
DB 404 AAAGAGCAGATTCCTCAAAACGAAATGAAGATGCTATCTCTCCATCCAGGAAA 463
QY 174 CGAAGTGGCAGTGGTTCGAGAGAGTGTGTACACTGTCACTTCAATATCA---TCCCCCA 230
DB 464 TGTGACACAGACCTACTCAGAGAGCTGTGTCTATACCTCTCAATCATCGGGTCTCTG 523
QY 231 TCAGAGATCTCTCCTCAGCTTCAATGATGCTGATGAGAACTTGAATGCTCTCCATCA 290
DB 524 TACAAGCCATCAGGAACTCTGCTGAGAGTACTATGAGATGTTCCCTGCAAGCTGA 583
QY 291 GAAAGTGAACAGTTAGAGAAAGGTCAGAGACAGAAATATGCCCTTCT---TAGGACTTC 347
DB 584 GAGACCCAGAGATCTCTGAGAGAACTGAGACTGAGTATTCATCTTACATATATGCCCTTC 643
QY 348 TGTAGTAGGCTTGTCTCTGACCCATGAGCATGATTAATGAGTGTGTGTTCCACAC 405
DB 644 TACAGACCCAGGAGTCCCGATCCCGAGAGATGAATGATGAATCTCTCATGCTCTCAC 701

RESULT 7
BQ389415
LOCUS 602414106F1 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:4522512 5',
mRNA sequence.
DEFINITION BQ389415.1 GI:13282861
ACCESSION BQ389415
VERSION BQ389415
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 924)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10423 row: l column: 01
High quality sequence stop: 726.
FEATURES
Location/Qualifiers
1..924
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4522512"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_92"
/notes="Organ: testis; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI. Cloned unidirectionally; oligo-dT primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
ORIGIN

```

```

Query Match      14.3%; Score 58; DB 12; Length 924;
Best Local Similarity 54.4%; Pred. No. 0.0015;
Matches 162; Conservative 0; Mismatches 130; Indels 6; Gaps 2;

```

QY	174	CGAGAAATGGCAGTGGTTCTGAAGAAAGTGTGTCTACACTGTTCATTAAATCACA---TCCCCCA	230
Db	403	TGTTGACAGACACCTACTCTCAGAGAGGTGTGTCTATACCTCTCATCATCATCGGGTCTCTG	462
QY	231	TCAGAGATCTCTCCTGAGTGTCCAATGATCATGGCTATGAGAACATTTGACTCCCTCACAAG	290
Db	463	TACAAGCCCATCAGGGAACCTCTGCTGAAGAGTACTACTGGAATGTTCCCTGCAAGCTGA	522
QY	291	GAAGTGTAGACAGTTTAGAGAAAGGTCTCAGACAGACAGNATATGCCCTTCT---TAGACATTC	347
Db	523	GAGACCCAGAGAGTCTTGGGAGGAACCTGAGACTCAGTATTTCACCTTCACATATGCGCTTC	582
QY	348	TGTTAGTAGGCGCTTGTTCTCTGCACCCATCAGCATGATTATGAAGTGTGTGTTCACAC	405
Db	583	TACAGACCCAGGCGATGCCCGATCCCCAGAGATGAATATGAACCTTCTCATGCGCTCAC	640
RESULT 9			
BM904106			
LOCUS			
DEFINITION	BM904106	1000 bp mRNA	linear EST 12-MAR-2002
ACCESSION	AGSCNCOURT_6692745	NIH MGC 85	Homo sapiens cDNA clone IMAGE:5499556
VERSION	5', mRNA sequence.		
KEYWORDS	BM904106		
SOURCE	BM904106.1	GI:193554351	
ORGANISM	Est.		
	Homo sapiens (human)		
	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
REFERENCE	1	(Bases 1 to 1000)	
AUTHORS	NIH-MGC http://mgi.nci.nih.gov/ .		
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-x@mail.nih.gov Tissue Procurement: Lou Staudt cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LHAM12133 row: b column: 05 High quality sequence stop: 698.		

```

FEATURES
source
1. 1000
/organism="Homo sapiens"
/mol_type="rRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5499556"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 85"
/notes="Organ: lymph; Vector: pCMV-SPORT6; Site: 1: NotI;
Site: 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.867 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH MGC Library."

```

114	AAACTTCAAGATCAAGTAAGAAGAAAGCCAGAAAGTTTCATCCATCTTAAATCAGGAAA	173
121		
128		
135		
142		
149		
156		
162		
169		
176	CGAGATGGCAGTGGTCTGAGAGAGTGGCTGCTACACTGTCTATTATGACA---TCCCCCA	230
183		
190		
197		
204		
211		
218		
225		
232		
239		
246		
253		
260		
267		
274	CGAGATGGCAGTGGTCTGAGAGAGTGGCTGCTACACTGTCTATTATGACA---TCCCCCA	330
281		
288		
295		
302		
309		
316		
323		
330		
337		
344	TGTTGACCACACCTACTTCGAGAGAGCTGTGCTATACCCCTCATCATCATCGGGTCTCTG	400
351		
358		
365		
372		
379		
386		
393		
400		

```

QY 231 TCAGAGATCCCTCCCTGAGCTCCAAATGATGCGCTATGAGAACATTTGACTCCCTCACAAG 290
DB 401 TACAAGGCCATCAGGAAGTCTGCTGAAGAGTACTATGAGAAATTTCCCTGCAAGCTGA 460
QY 291 GAAAGTGACAGATTTAGAGAAAGTTCAGACAGATATATCCCTTCT---TAGGACTTC 347
DB 461 GAGACCCAGAGAGTCTTGGGAGGAATCTGAGACTGAGTATTCATCTTACATATATGCTTC 520
QY 348 TGTTAGTAGGCTTCTCTCCGACCCATGAGCATGATGATGAAGTGTGTTTCCACAC 405
DB 521 TACAGACCCAGGCGATGCCCGATCCCGAAGATGATGAACTTTCTCATGCTTCAC 578

RESULT 10
BM479887 1039 bp mRNA linear EST 05-FEB-2002
LOCUS AGENCOURT_6464830 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:5577364
DEFINITION 5', mRNA sequence.
ACCESSION BM479887
VERSION BM479887.1 GI:18528929
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1039)
TITLE NIH-MGC http://mgi.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLMI2330 row: 1 column: 05
High quality sequence stop: 665.
Location/Qualifiers
1. .1039
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5577364"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 92"
/notes="Organ: testis; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

ORIGIN
Query Match 14.3%; Score 58; DB 12; Length 1039;
Best Local Similarity 54.4%; Pred. No. 0.0018;
Matches 162; Conservative 0; Mismatches 130; Indels 6; Gaps 2;

QY 114 AAACTTCAAGATCAAGATGAAGAACCCAGAGTTTCATCCACTTCTATCAGGAAA 173
DB 245 AAAGAGGCAAGATTCCTCAAAACGAAATGAAGAATGTCTATCTCCCATCCAGGAAA 304
QY 174 CGAGATGGCAGTGGTTCGAGAAAGTGTCTACACTGTCTATTAATCACA---TCCCCCA 230
DB 305 TGTGACCAAGCTTCTCAGAGAGCTGTGCTATACCTCATCATCATCGGGTTCTG 364
QY 231 TCAGAGATCCCTGAGCTCCAAATGATGCTATGAGAACATTTGACTCCCTCACAAG 290
DB 365 TACAAGGCCATCAGGGAATCTGCTGAAGAGTACTATGAGAAATGTTCCTGCAAGCTGA 424
QY 291 GAAAGTGAGACAGTCTTGAAGAAAGTTCAGACAGATATGCGCTTCT---TAGGACTTC 347

FEATURES
source
1. .1070
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5496261"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 85"
/notes="Organ: lymph; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.867 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

ORIGIN
Query Match 14.3%; Score 58; DB 12; Length 1070;
Best Local Similarity 54.4%; Pred. No. 0.0018;
Matches 162; Conservative 0; Mismatches 130; Indels 6; Gaps 2;

QY 114 AAACTTCAAGATCAAGATGAAGAACCCAGAGTTTCATCCACTTCTATCAGGAAA 173
DB 306 AAAGAGGCAAGATTCCTCAAAACGAAATGAAGAATGTCTATCTCCCATCCAGGAAA 365
QY 174 CGAGATGGCAGTGGTTCGAGAAAGTGTCTACACTGTCTATTAATCACA---TCCCCCA 230
DB 366 TGTGACCAAGCTTCTCAGAGAGCTGTGCTATACCTCATCATCATCGGGTTCTG 425
QY 231 TCAGAGATCCCTGAGCTCCAAATGATGCTATGAGAACATTTGACTCCCTCACAAG 290
DB 426 TACAAGGCCATCAGGGAATCTGCTGAAGAGTACTATGAGAAATGTTCCTGCAAGCTGA 485
QY 291 GAAAGTGAGACAGTCTTGAAGAAAGTTCAGACAGATATGCGCTTCT---TAGGACTTC 347
DB 486 GAGACCCAGAGAGTCTTGGGAGGAATCTGAGACTGAGTATTCATCTTACATATGCTTC 545
QY 348 TGTTAGTAGGCTTCTCTCCGACCCATGAGCATGATGATGAAGTGTGTTTCCACAC 405

```

```

DB 425 GAGACCCAGAGAGTCTTGGGAGGAAGTGAAGTCTGAGTATTCACATATGCTTC 484
QY 348 TGTTAGTAGGCTTCTCTCCGACCCATGAGCATGATGATGAAGTGTGTTTCCACAC 405
DB 485 TACAGACCCAGGCGATGCCCGATCCCGAAGATGAATGAACTTCTCATGCTTCAC 542

RESULT 11
BM456595 1070 bp mRNA linear EST 05-FEB-2002
LOCUS AGENCOURT_6409034 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:5496261
DEFINITION 5', mRNA sequence.
ACCESSION BM456595
VERSION BM456595.1 GI:18505635
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1070)
TITLE NIH-MGC http://mgi.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Lou Staudt
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLMI2124 row: h column: 22
High quality sequence stop: 648.
Location/Qualifiers
1. .1070
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5496261"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 85"
/notes="Organ: lymph; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.867 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

ORIGIN
Query Match 14.3%; Score 58; DB 12; Length 1070;
Best Local Similarity 54.4%; Pred. No. 0.0018;
Matches 162; Conservative 0; Mismatches 130; Indels 6; Gaps 2;

QY 114 AAACTTCAAGATCAAGATGAAGAACCCAGAGTTTCATCCACTTCTATCAGGAAA 173
DB 306 AAAGAGGCAAGATTCCTCAAAACGAAATGAAGAATGTCTATCTCCCATCCAGGAAA 365
QY 174 CGAGATGGCAGTGGTTCGAGAAAGTGTCTACACTGTCTATTAATCACA---TCCCCCA 230
DB 366 TGTGACCAAGCTTCTCAGAGAGCTGTGCTATACCTCATCATCATCGGGTTCTG 425
QY 231 TCAGAGATCCCTGAGCTCCAAATGATGCTATGAGAACATTTGACTCCCTCACAAG 290
DB 426 TACAAGGCCATCAGGGAATCTGCTGAAGAGTACTATGAGAAATGTTCCTGCAAGCTGA 485
QY 291 GAAAGTGAGACAGTCTTGAAGAAAGTTCAGACAGATATGCGCTTCT---TAGGACTTC 347
DB 486 GAGACCCAGAGAGTCTTGGGAGGAATCTGAGACTGAGTATTCATCTTACATATGCTTC 545
QY 348 TGTTAGTAGGCTTCTCTCCGACCCATGAGCATGATGATGAAGTGTGTTTCCACAC 405

```

Db 546 TACAGACCCAGGCGATCCCGATCCCGAGAGATGAATATGAATCTTCATGCTCAC 603

RESULT 12
BG106563 602290373F1 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:4385178 5',
LOCUS mRNA sequence.
DEFINITION EST. 1095 bp mRNA linear EST 30-JAN-2001

ACCESSION BG106563
VERSION BG106563
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Louis Staudt, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL0066 row: a column: 19
High quality sequence stop: 641.
Location/Qualifiers
1..1095
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4385178"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_85"
/note="Organ: lymph; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.867 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

FEATURES
source
1..1095
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4385178"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_85"
/note="Organ: lymph; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.867 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

ORIGIN
Query Match 14.3%; Score 58; DB 12; Length 1095;
Best Local Similarity 54.4%; Pred. No. 0.0018;
Matches 162; Conservative 0; Mismatches 130; Indels 6; Gaps 2;
QY 114 AAACTTCAAGATCAAGATGAAGAAAGCCAGAGTTTCATCCACTTCTAATCAGGAAA 173
Db 207 AAAGAGGCAAGATTCCCAAAACGAAATGAAGAATGTCTACTTCCCATCCAGGCAA 266
QY 174 CGAGATGGCAGTGGTTCTGAAGAAGTGTGCTACACTGTCAATTAATCACA---TCCCCCA 230
Db 267 TGTGACCAAGCTTCTAGAGAGCTGTGCTATACCTCATCATCGGGTCTCTG 326
QY 231 TCAGAGATCCTCCCTGAGTCCCAATGATGCTGCTATGAGACATTCAGTCCCTCACAAG 290
Db 327 TACAAGGCCATCAGGGAATCTGCTGAAGAGTACTATGAGAAATGTTCCCTGCAAGAGCTGA 386
QY 291 GAAAGTGACAGCTTTAGAGAAAGGTCAGAGACAGAAATATGCCCTTCT---TAGGACTTC 347
Db 387 GAGACCCAGAGCTCCTTGGAGGAATCTGAGCTGAGTAATTCATTCTACATATGCCCTTC 446
QY 348 TGTAGTAGGCTTGTTCCTGCAACCCATGAGCATGATTAAGAGTTGTGTTTCCACAC 405
Db 447 TACAGACCCAGGCGATCCCGATCCCGAGAGATGAATATGAATCTTCATGCTCAC 504

RESULT 13
BF795688

LOCUS BF795688 1124 bp mRNA linear EST 12-JAN-2001
DEFINITION 602259560F1 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:4342719 5',
mRNA sequence.
ACCESSION BF795688
VERSION BF795688
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Louis Staudt, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL9958 row: h column: 16
High quality sequence stop: 715.
Location/Qualifiers
1..1124
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4342719"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_85"
/note="Organ: lymph; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.867 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

ORIGIN
Query Match 14.3%; Score 58; DB 10; Length 1124;
Best Local Similarity 54.4%; Pred. No. 0.0018;
Matches 162; Conservative 0; Mismatches 130; Indels 6; Gaps 2;
QY 114 AAACTTCAAGATCAAGATGAAGAAAGCCAGAGTTTCATCCACTTCTAATCAGGAAA 173
Db 282 AAAGAGGCAAGATTCCCAAAACGAAATGAAGAATGTCTACTTCCCATCCAGGCAA 341
QY 174 CGAGATGGCAGTGGTTCTGAAGAAGTGTGCTACACTGTCAATTAATCACA---TCCCCCA 230
Db 342 TGTGACCAAGCTTCTAGAGAGCTGTGCTATACCTCATCAATCATCGGGTCTCTG 401
QY 231 TCAGAGATCCTCCCTGAGTCCCAATGATGCTGCTATGAGACATTCAGTCCCTCACAAG 290
Db 402 TACAAGGCCATCAGGGAATCTGCTGAAGAGTACTATGAGAAATGTTCCCTGCAAGAGCTGA 461
QY 291 GAAAGTGACAGCTTTAGAGAAAGGTCAGAGACAGAAATATGCCCTTCT---TAGGACTTC 347
Db 462 GAGACCCAGAGCTCCTTGGAGGAATCTGAGCTGAGTAATTCATTCTACATATGCCCTTC 521
QY 348 TGTAGTAGGCTTGTTCCTGCAACCCATGAGCATGATTAAGAGTTGTGTTTCCACAC 405
Db 522 TACAGACCCAGGCGATCCCGATCCCGAGAGATGAATATGAATCTTCATGCTCAC 579

RESULT 14
BM455198 1193 bp mRNA linear EST 05-FEB-2002
LOCUS AGENCOURT 6405645 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:5500141
DEFINITION 5', mRNA sequence.
ACCESSION BM455198
VERSION BM455198.1 GI:18504238

KEYWORDS EST. Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1193)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue procurement: Lou Staudt
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC Clone Distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12134 row: j column: 14
High quality sequence stop: 592.

FEATURES
source
1..1193
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:550141"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_85"
/note="Organ: lymph; Vector: pCMV-Sport6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.867 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

ORIGIN
Query Match 14.3%; Score 58; DB 12; Length 1193;
Best Local Similarity 54.4%; Pred. No. 0.0018;
Matches 162; Conservative 0; Mismatches 130; Indels 6; Gaps 2;
QY 114 AAAATCTCAAGATCAAGATAAGAAAGCAAGAGTTTCATCCATCTTCAATCAGGAAA 173
Db 235 AAAGAGGCAAGATCCCAAAAGAAATGAAGATGATCATCTCCATCCAGGACAA 294
QY 174 CGAGATGCGACGTGGTCTTGAAGAAGTGTGCTACACTGTCATTAATACAA---TCGCCCA 230
Db 295 TGTGACACAGACTACTCAGAGAGGTGTGCTATACCTCATCATCATCGGTTCTCTG 354
QY 231 TCAGAGATCTCCCTGAGCTCCAATGATGCTGCTATGAGACATTCCTCCTCACAAG 290
Db 355 TACAAGGCCATCAGGGAATCTGCTGAAGAGTACTATGAGATGTTCCCTGCAAGCTGA 414
QY 291 GAAAGTGAGACAGTTTAGAAGAAGGTTCAGAGACAGATATGCCCTTCT---TAGGACTTC 347
Db 415 GAGACCCAGAGAGTCTCTGGAGGAACCTGAGACTGAGTATTCATCTTACATATGCCTTC 474
QY 348 TGTAGTAGGCTGTCTCTGACCATGATGATGATGATGATGATGATGATGATGATGATG 405
Db 475 TACAGACCCAGGCGATGCCCGATCCCAAGATGAATGAATGAATGAATGAATGAATGAAT 532

RESULT 15
CB286394
LOCUS
DEFINITION
CMD36_D01_13 UNMMP3 Sus scrofa cDNA clone PPSUBLIB_21D01 5', mRNA
sequence.
ACCESSION
CB286394
VERSION
CB286394.1 GI:28577348
KEYWORDS
EST.
SOURCE
Sus scrofa (pig)
ORGANISM
Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

REFERENCE 1 (bases 1 to 526)
Dvorak, C.M.T., Hyland, K.A., Zhang, Y., Fahrenkrug, S.C. and
Murtaugh, M.P.
Porcine jejunal Peyer's patch expressed sequences
Unpublished (2003)
Contact: Murtaugh, M.P.
Department of Veterinary Pathobiology
University of Minnesota
1971 Commonwealth Ave., St. Paul, MN 55108, USA
Tel: 612-625-6735
Fax: 612-625-5203
Email: murta001@umn.edu

The PHRED quality scores of this sequence are supplied below: 51 39
39 39 40 35 28 28 25 25 28 35 35 35 35 35 35 35 39 39 39 40
40 35 35 35 35 35 35 45 51 51 56 56 56 40 40 35 35 35 35 35
56 56 56 51 51 45 28 22 15 15 28 30 40 40 35 35 35 56 56 56
45 45 51 56 51 51 45 43 43 43 43 43 43 45 43 43 43 43 43 40
43 36 36 36 36 36 43 43 43 43 43 43 43 43 43 43 43 43 43 43
45 43 43 43 43 43 43 43 43 43 43 43 43 43 43 43 43 43 43 43
51 51 45 41 41 51 51 45 43 43 43 43 43 43 43 43 43 43 43 43
43 35 43 43 43 43 43 43 43 43 43 43 43 43 43 43 43 43 43 43
45 45 45 45 45 45 40 40 40 40 40 40 40 40 40 40 40 40 40 40
56 56 51 51 51 51 56 56 56 51 51 51 46 46 56 56 56 40 40 40 40
39 48 56 56 56 51 51 51 51 46 46 56 56 56 56 56 56 56 40 40 40 40
51 51 51 45 46 46 46 46 46 46 56 56 56 56 56 56 56 56 56 56
56 56 56 51 51 45 45 45 45 45 45 51 51 51 51 51 51 51 51 40 40
37 35 39 37 40 40 40 40 40 40 40 40 40 40 40 40 40 40 40 40
33 32 32 22 32 32 32 32 32 32 36 40 40 40 40 40 40 40 40 40
46 35 35 35 35 35 35 46 51 42 35 37 35 35 35 35 35 35 35 35
45 45 40 40 34 34 34 34 36 39 46 56 48 37 34 32 32 32 25 21
18 23 40 32 32 34 34 36 39 46 56 48 37 34 32 32 32 32 29 25
34 37 40 40 40 40 40 40 40 40 40 40 40 40 40 40 40 40 40 40
26 25 23 18 23 25 23 25 23 19 25 48 40 29 29 29 29 29 29 27 23
47 47 56 42 42 42 42 42 42 42 42 40 40 40 40 40 40 40 40 29 24
25 32 34 40 40 30 32 26 26 25 29 29 25 28 32 32 29 Sequences
were generated on ABI 3100 and 3700 capillary sequencers. Raw
sequence data was batch processed analyzed, groomed and submitted
by the Animal Biotechnology Center (URL:
http://primer.ansci.umn.edu/software.html). University of
Minnesota. Trace files have been submitted to the NCBI trace
archive (http://www.ncbi.nlm.nih.gov/Traces/trace.cgi?).

Chromatograms were analyzed with Phred Version: 0.000925.c . Vector
(pCMV-Sport6, Invitrogen), Sus scrofa mitochondrion
(gi|5835862|ref NC_000845.1), Escherichia coli K12 (gi|5935862|
ref NC_000913.1), Porcine reproductive and respiratory syndrome
virus (gi|11878202|gb|AF303357.1|AF303357), Alteromonas sp.16S rRNA
gene (gi|4218471|emb|Y15322.1|Y15322), Sus scrofa 28S ribosomal
RNA (gi|3388170|gb|AF080393.1|AF080393), and Sus scrofa 18S ribosomal
RNA (gi|3985982|gb|AF102857.1|AF102857). Sequences were identified
and masked by crossmatch -mismatch=20 -minmatch=12, penalty=-8.
SWAT alignment scores were set to match score=1, gap initial
penalty=-3 and gap extension penalty=-2. A continuous stretch of at
least 50 bases of a PHRED quality 20 were required for
submissions 5' and 3' TERMINI were defined by a continuous stretch
of 10 nucleotides with a PHRED quality not less than 20. A terminal
stretch of at most 40 un-masked bases were automatically trimmed
when flanking masked sequence. Sequences with an INTERNAL continuous
stretch of at most 20 bases with PHRED quality less than 20 were
automatically prepared for submission. HIGHQUAL_START and HIGHQUAL_STOP
refer to the coordinates on CLIP_QUALITY and CLIP_VECTOR at the
NCBI trace archive.
Plate: 21 row: D column: 01
Seq primer: SP6 primer
High quality sequence start: 67
High quality sequence stop: 334.
Location/Qualifiers
1..526
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/clone="PPSUBLIB_21D01"

FEATURES
source

/tissue type="Peyer's patch"
/clone lib="UMNMF3"
/notes=Organ: small intestine, jejunum; Vector:
pCMVSPORT6; Site 1: NotI; Site 2: SalI; Jejunal Peyer's
patches were isolated from healthy, 4-6 week old
cross-bred pigs. RNA was extracted either immediately or
after 3 hours stimulation in an Ussing chamber with one of
the following treatments: 1) no treatment, 2) Salmonella
choleraesuis vaccine strain SC-54, 3) phorbol myristate
acetate, concanavalin A, and 8-bromo-cyclic AMP, and 4)
lipopolysaccharide and cholera toxin. Each treatment was
performed in the presence and absence of cycloheximide.
Purified poly A+ RNA from each of the treated tissues
(2-4) was combined together, reverse transcribed, and
cloned in to pCMVSPORT6 to make a library of approximately
530,000 recombinant clones with an average insert size of
1.0 kbp. Poly A+ RNA from freshly isolated and non-treated
tissue from an Ussing chamber (treatment 1) was cloned in
the same manner to produce an unstimulated cDNA library of
approximately 900,000 clones with an average insert size
of 1.5 kbp. Equal portions of the two libraries were
pooled and then subtracted with porcine ST fibroblast RNA
to create a subtracted porcine Peyer's patch library of
approximately 6,000 clones with an average insert size of
1.0 kbp."

ORIGIN

Query Match	14.2%;	Score 57.4;	DB 14;	Length 526;
Best Local Similarity	51.8%;	Pred. No. 0.0024;		
Matches 155;	Conservative 0;	Mismatches 141;	Indels 3;	Gaps 1;

Qy	106	TTTGAAGAAAACCTTCAGATCAAGATAAGAAAGCCCAAGAGTTTCATCCACTTCTAAT	165
Db	2	TTTAAAGCAACACAGATTCCTCCCAAGCAAAATGAGAAACATCATCTTCTCCCTCCAG	61

Qy	166	CAGGAAAACGAGATGGCAGTGGTTCTGAGAGAGTGCTGCTACACTGTCATTATCACATC	225
Db	62	GAGGAGATGTTGACCAGAGTCCGCTGGAGGATCTGTCTACACCCCTCATTGATCACAGC	121

Qy	226	CCCCATCAGATCCTCCTGAGTCCCAATGATGATG---GCTATGAGAACATTGACTCC	282
Db	122	ACCCCTGGAGAGGCCAATAGGCACCCCTGCTGAGGAGTGTCTATGAGACGTTTCCCTC	181

Qy	283	CTCACAGGAAAGTGAGACAGTTTAGAGAAAGGTGAGAGACAGAAATATGCCCTTCTTAGG	342
Db	182	AAGACCGAGAGGCCCCAGAGAGTCAATTGGAGGAAACGAGAGTACTCACTTATTTCAT	241

Qy	343	ACTTCTGTAGTAGCCCTTGTTCTGACCCATGACATGATTATGAGTGTGTTTCC	401
Db	242	GTGTCTTCTACTCTTGGCCAGCCATCCTCCTCAGAAATGAATGAACTTCTCATGCC	300

Search completed: April 22, 2004, 16:07:50
Job time : 2542 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 23, 2004, 03:10:01 ; Search time 14153 Seconds
(without alignments)
11520.985 Million cell updates/sec

Title: US-10-005-907-1
Perfect score: 3762
Sequence: 1 ggaacacgagtcactgtga.....aaaaaaaaaaaaaaaaaaaaa 3762

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 3470272 seqs, 21671516995 residues

Word size : 0
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 150 summaries

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.btg.*
- 3: gb.in.*
- 4: gb.cm.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg.hum.*
- 31: em.htg.inv.*
- 32: em.htg.other.*
- 33: em.htg.mus.*
- 34: em.htg.pln.*
- 35: em.htg.rod.*
- 36: em.htg.mam.*
- 37: em.htg.vrt.*
- 38: em.sv.*
- 39: em.htgo.hum.*
- 40: em.htgo_mus.*
- 41: em.htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3762	100.0	3762	6	AX505122 Sequence
2	3359	89.3	185467	9	AL606804 Human DNA
3	3284	87.3	4032	9	AK126682 Homo sapi
4	3144	83.6	4136	9	AK124520 Homo sapi
5	3016	80.2	141268	2	AC074365 Homo sapi
6	2506	65.6	2648	9	BC024174 Homo sapi
7	1624	43.2	1709	9	AK123798 Homo sapi
8	681	18.1	121353	2	AC023386 Homo sapi
9	479	12.7	121353	2	AC023386 Homo sapi
10	106	2.8	1924	9	AB060908 Macaca fa
11	62	1.6	144333	9	AC105036 Homo sapi
12	52	1.4	217	9	HS0805906
13	52	1.4	784	5	BC052102 Xenopus l
14	52	1.4	1815	9	BC009760 Homo sapi
15	50	1.3	570	9	HS0807892
16	50	1.3	1166	9	BC043545 Homo sapi
17	50	1.3	1581	9	BC034284 Homo sapi
18	50	1.3	1589	9	BC008899 Homo sapi
19	50	1.3	1616	10	BC043714 Mus muscu
20	50	1.3	1639	5	BC056324 Danio rer
21	50	1.3	1654	9	BC040896 Homo sapi
22	50	1.3	1737	3	AY069228 Drosophil
23	50	1.3	1786	5	BC053205 Danio rer
24	50	1.3	1820	9	AK026959 Homo sapi
25	50	1.3	1880	5	BC052972 Danio rer
26	50	1.3	2275	9	BC036869 Homo sapi
27	50	1.3	2481	9	AB052200 Macaca fa
28	50	1.3	3606	9	AY078404 Homo sapi
29	50	1.3	96367	9	AC127894 Homo sapi
30	49	1.3	159	6	AR425685 Sequence
31	49	1.3	159	6	BD121238 EST and e
32	49	1.3	444	9	BC038545 Homo sapi
33	49	1.3	476	6	AX575591 Sequence
34	49	1.3	704	9	BC063013 Homo sapi
35	49	1.3	752	10	BC060283 Mus muscu
36	49	1.3	959	9	S71381 prosome bet
37	49	1.3	997	9	BC032451 Homo sapi
38	49	1.3	1008	9	BC000761 Homo sapi
39	49	1.3	1030	9	BC029349 Homo sapi
40	49	1.3	1051	9	BC038410 Homo sapi
41	49	1.3	1068	9	BC004905 Homo sapi
42	49	1.3	1109	10	BC063150 Rattus no
43	49	1.3	1154	9	BC040307 Homo sapi
44	49	1.3	1320	5	BC056295 Danio rer
45	49	1.3	1346	10	BC032932 Mus muscu
46	49	1.3	1429	6	AX587564 Sequence
47	49	1.3	1429	3	AF097996 Homo sapi
48	49	1.3	1440	3	AY297438 Aiptasia
49	49	1.3	1499	9	BC042195 Homo sapi
50	49	1.3	1595	9	BC008836 Homo sapi
51	49	1.3	1655	9	BC039725 Homo sapi
52	49	1.3	1661	9	BC025753 Homo sapi
53	49	1.3	1676	6	AX188231 Sequence
54	49	1.3	1687	10	BC035953 Mus muscu
55	49	1.3	1706	3	AK112205 Ciona int
56	49	1.3	1729	9	BC012854 Homo sapi
57	49	1.3	1737	10	BC061980 Rattus no
58	49	1.3	1786	9	AB070107 Macaca fa
59	49	1.3	1806	3	AY118452 Drosophil
60	49	1.3	1835	9	BC029874 Homo sapi
61	49	1.3	1838	5	BC061320 Silurana
62	49	1.3	1850	9	BC044938 Homo sapi
63	49	1.3	1894	9	AB063046 Macaca fa
64	49	1.3	1933	6	AR300635 Sequence
65	49	1.3	1933	6	AX146886 Sequence

301 GACTCCCTCAAGGAAAGTGAGACAGTTTAGAGAAAGGTGACAGACAGAAATATGCCCC 360
Db GACTCCCTCAAGGAAAGTGAGACAGTTTAGAGAAAGGTGACAGACAGAAATATGCCCC 360
361 CTTAGGACTTCTGTAGTGGCCCTGTTCTGTCACCCATGAGCATGATATGAGTTG 420
Db CTTAGGACTTCTGTAGTGGCCCTGTTCTGTCACCCATGAGCATGATATGAGTTG 420
421 TTTCCACACATAAATCCTCAAGCTGCTTTATCATCCTTCCAGCAATGAAGCAATGAGAA 480
Db TTTCCACACATAAATCCTCAAGCTGCTTTATCATCCTTCCAGCAATGAAGCAATGAGAA 480
481 TAGCAGACTCTGGGAAAGTTGTCACCTGAGCAGTGCATGAACATCTCTTCTGGCTA 540
Db TAGCAGACTCTGGGAAAGTTGTTCACTGAGCAGTGCATGAACATCTCTTCTGGCTA 540
541 AAGTTTAGAAATATATCTTATATATATCTTATGAGCACTCTGATATGAGCATCTCTG 600
Db AAGTTTAGAAATATATCTTATATATATCTTATGAGCACTCTGATATGAGCATCTCTG 600
601 TGGCTTAGGTAATCATAGAAATGACACATGACCTTAAATATCTTATGCTGTTTTC 660
Db TGGCTTAGGTAATCATAGAAATGACACATGACCTTAAATATCTTATGCTGTTTTC 660
661 TTGTAAAGTTTGGAGACATGAGGTGATATAAATACTTTCTAGGCAATATGTAATA 720
Db TTGTAAAGTTTGGAGACATGAGGTGATATAAATACTTTCTAGGCAATATGTAATA 720
721 TGAATAAATTTCTAATCCCTGACTAATGAGTGAACCTCTTCTAGGCCAAGAGA 780
Db TGAATAAATTTCTAATCCCTGACTAATGAGTGAACCTCTTCTAGGCCAAGAGA 780
781 CCTCAGATGAACTGAAAGCTGAAATCTGCGCATGATAGGAGGAGGTGAGACACACC 840
Db CCTCAGATGAACTGAAAGCTGAAATCTGCGCATGATAGGAGGAGGTGAGACACACC 840
841 TTGTTATACCCCTTCCCTTTTGGAGTTATGCAAGTGAACCAAGTGAAGTGAAGACT 900
Db TTGTTATACCCCTTCCCTTTTGGAGTTATGCAAGTGAACCAAGTGAAGTGAAGACT 900
901 GATGAATAGACTGATGTCGCAATAGAGTCCCAATTCACCACTGACTCTGGGTAGAT 960
Db GATGAATAGACTGATGTCGCAATAGAGTCCCAATTCACCACTGACTCTGGGTAGAT 960
961 CACACACTGTCTGAGGAAATCCATCTATGAGACTTTGTCTACATAACAGAGACCTTGGT 1020
Db CACACACTGTCTGAGGAAATCCATCTATGAGACTTTGTCTACATAACAGAGACCTTGGT 1020
1021 TCCACACCCCTTATTTTAGTAAAGCATCTTTCTACTGACTCTTAACTTTTGA 1080
Db TCCACACCCCTTATTTTAGTAAAGCATCTTTCTACTGACTCTTAACTTTTGA 1080
1081 CAAGCTTAACCTCTTCAACCAATGCAATCAGACAACTTTGAATCTACTATGACCT 1140
Db CAAGCTTAACCTCTTCAACCAATGCAATCAGACAACTTTGAATCTACTATGACCT 1140
1141 GTAAGCTCTCTGCTTCAAGATCTTGGCTTTTAACTGAAACCGATGCACTTTCCA 1200
Db GTAAGCTCTCTGCTTCAAGATCTTGGCTTTTAACTGAAACCGATGCACTTTCCA 1200
1201 TTTAATGATTTATGCTTTGCTTGAATCTCTGCTCTCTTAAATGATATAAAGTAAACG 1260
Db TTTAATGATTTATGCTTTGCTTGAATCTCTGCTCTCTTAAATGATATAAAGTAAACG 1260
1261 GTGACCTGACCACTCAGGACACTTTCTAGGACCTCTCTGAGAGTGTATCCAGGCCAT 1320
Db GTGACCTGACCACTCAGGACACTTTCTAGGACCTCTCTGAGAGTGTATCCAGGCCAT 1320
1321 GGTAGTCACTGCTGCTCAGATCACTCTTAAATATTTACAGATTTGGTTTGG 1380
Db GGTAGTCACTGCTGCTCAGATCACTCTTAAATATTTACAGATTTGGTTTGG 1380

1381 TTACCAATAGCTCTCCACAAATATATGTCGAAGAACTTTCAATTCGAAGCTGCTCACCA 1440
Db TTACCAATAGCTCTCCACAAATATATGTCGAAGAACTTTCAATTCGAAGCTGCTCACCA 1440
1441 AATTTCAAAATGCCAACATCTCCCATCCAAATTAATCTATTTTCACTTTGAGGTGAATCTA 1500
Db AATTTCAAAATGCCAACATCTCCCATCCAAATTAATCTATTTTCACTTTGAGGTGAATCTA 1500
1501 CTCAATAAATCTGTGTAGACCAAGTGCAGACCCCTTGTCTAACCCTGACATTTACTTCAAT 1560
Db CTCAATAAATCTGTGTAGACCAAGTGCAGACCCCTTGTCTAACCCTGACATTTACTTCAAT 1560
1561 TTTTCTTTTCTATCTACTGATATTTTGTGATATAAATTTTGCATATAAATCTTGAGTAAAT 1620
Db TTTTCTTTTCTATCTACTGATATTTTGTGATATAAATTTTGCATATAAATCTTGAGTAAAT 1620
1621 TAATAGTTTTTGCATCTGCTTTTCTGAGAGAGAAATGAAAGTGTCAAAATATAAAA 1680
Db TAATAGTTTTTGCATCTGCTTTTCTGAGAGAGAAATGAAAGTGTCAAAATATAAAA 1680
1681 AAGATGAATGAAGCATATATAAATTTTCAATTTTCTAGTCAACAGAGAAATC 1740
Db AAGATGAATGAAGCATATATAAATTTTCAATTTTCTAGTCAACAGAGAAATC 1740
1741 GAAGGATCTGTTCAAAATATTAAGTAAATTAAGTAAATTAAGTAAATTAAGTAAAT 1800
Db GAAGGATCTGTTCAAAATATTAAGTAAATTAAGTAAATTAAGTAAATTAAGTAAAT 1800
1801 GCAACACACTAGTAAATTTAACTGTGACTAGTAAATCTCCGAGGTGAGTGTAGT 1860
Db GCAACACACTAGTAAATTTAACTGTGACTAGTAAATCTCCGAGGTGAGTGTAGT 1860
1861 TTCTGTTTTTAAATTTCAAGCAAACTGGAAATAATCCATTAATTTCTTTTCTTCC 1920
Db TTCTGTTTTTAAATTTCAAGCAAACTGGAAATAATCCATTAATTTCTTTTCTTCC 1920
1921 AAGAGTTTTTAAATGATATGCGAGCTTCTAATTTGGAGCAAAAGCTTAAATGACAA 1980
Db AAGAGTTTTTAAATGATATGCGAGCTTCTAATTTGGAGCAAAAGCTTAAATGACAA 1980
1981 TGCAATTCATTAATATTTTTTGTATAGTTACAGTATACGAGTTCAGTATCCCTTAGATG 2040
Db TGCAATTCATTAATATTTTTTGTATAGTTACAGTATACGAGTTCAGTATCCCTTAGATG 2040
2041 AGATGCTGGGACCAAGAGTGTGGATTCAGATTTATTTTGGATTTTGGATTTT 2100
Db AGATGCTGGGACCAAGAGTGTGGATTCAGATTTATTTTGGATTTTGGATTTT 2100
2101 CCATACATATAATGAGAGAGTTGGAAATGGGATTCAGTCTAATCAATAAAATTCACCTA 2160
Db CCATACATATAATGAGAGAGTTGGAAATGGGATTCAGTCTAATCAATAAAATTCACCTA 2160
2161 TGTTGATATACACCTTATCTGAATAGCTGAGGTAAATTTTATACAAATTTTAAATAA 2220
Db TGTTGATATACACCTTATCTGAATAGCTGAGGTAAATTTTATACAAATTTTAAATAA 2220
2221 TTTTATGCTTGAACACAGAGTTTGGCACAATTCGACATTCGACATTCAGTAT 2280
Db TTTTATGCTTGAACACAGAGTTTGGCACAATTCGACATTCGACATTCAGTAT 2280
2281 TTCAGTCACTGCTCAAAAAGTTTTCAGATGTTTAAAGCTGATGAGTTCATGCCAGTGA 2340
Db TTCAGTCACTGCTCAAAAAGTTTTCAGATGTTTAAAGCTGATGAGTTCATGCCAGTGA 2340
2341 TCCGAGTCTTTTGGGAGCCACAGAGTGTGATCTCTGAGCCCGAGGCTTTGAGGCCAG 2400
Db TCCGAGTCTTTTGGGAGCCACAGAGTGTGATCTCTGAGCCCGAGGCTTTGAGGCCAG 2400
2401 ACTGCAACACAGTGAACCTGCTTTTCAAAATTTTAAATAATTTAGCCAGGTGTGT 2460
Db ACTGCAACACAGTGAACCTGCTTTTCAAAATTTTAAATAATTTAGCCAGGTGTGT 2460
2461 GGTGACACCTGTAGTCCCAAGGTACTCAGAGGCTGAGGTAGGTAGTGTGAGACTG 2520

2461	Db	GGTGACACACTGTGTAGTCCCGAGGTACTCAGGAGGCTGAGGTAGTAGGATTTGTGAGACTG	2520
2521	Qy	GGAGGTTTGAGGCTGAATCAGCCAGGATCTTGCCACCAATCCAGCTTTGGGCAACAGAG	2580
2521	Db	GGAGGTTTGAGGCTGAATCAGCCAGGATCTTGCCACCAATCCAGCTTTGGGCAACAGAG	2580
2581	Qy	TGAGACCCGTGTCTCAAAAAAAGAGTTTCAGATTTTGAGAGCATTTTCAGATCTTCAG	2640
2581	Db	TGAGACCCGTGTCTCAAAAAAAGAGAGTTTCAGATTTTGAGAGCATTTTCAGATCTTCAG	2640
2641	Qy	ATTAGGAGATTTTCAACCTGTACTGACCTTTAGTCAATTCAGACGATTAATCAATAGGTG	2700
2641	Db	ATTAGGAGATTTTCAACCTGTACTGACCTTTTAGTCAATTCAGACGATTAATCAATAGGTG	2700
2701	Qy	GACTCCAGATAACTCAATTTGCTGTATACCAATTTTGCTCTCTPATTTCAACGAATCTTAT	2760
2701	Db	GACTCCAGATAACTCAATTTGCTGTATACCAATTTTGCTCTCTPATTTCAACGAATCTTAT	2760
2761	Qy	GCCTCTTGTGGTGATTTAATCTGCGGAGGAAACAATAGAAATTTTGCAATTTCTAGA	2820
2761	Db	GCCTCTTGTGGTGATTTAATCTGCGGAGGAGAAACAATAGAAATTTTGCAATTTCTAGA	2820
2821	Qy	AAAGTCAATCTGTCAAAATATGTCAGTCTCTGTAGATATTAGCCAAATTTTAGGAAAATGAC	2880
2821	Db	AAAGTCAATCTGTCAAAATATGTCAGTCTCTGTAGATATTAGCCAAATTTTAGGAAAATGAC	2880
2881	Qy	AAAATTTTTTACTTTTTCGTCTGCTTTGTAGCTGTTTTATGATATAAATACCTTATTTGT	2940
2881	Db	AAAATTTTTTACTTTTTCGTCTGCTTTGTAGCTGTTTTATGATATAAATACCTTATTTGT	2940
2941	Qy	AATAAAATTAATTTTAATTTTGTGTAAACAATCTGGAATTTATCAGAGAAGGGCAAGCAATA	3000
2941	Db	AATAAAATTAATTTTAATTTTGTGTAAACAATCTGGAATTTATCAGAGAAGGGCAAGCAATA	3000
3001	Qy	GGTTAAATAAACAGTATTTGTTGTAGAGAAAGAAAGTTGAAATCCAGAGCAATCAATGCTTT	3060
3001	Db	GGTTAAATAAACAGTATTTGTTGTAGAGAAAGAAAGTTGAAATCCAGAGCAATCAATGCTTT	3060
3061	Qy	CTGTGTCTCACCATAGCCACAGCAGATGCTTTAATCTTTCCGAGATCTAGTTTTTCAG	3120
3061	Db	CTGTGTCTCACCATAGCCACAGCAGATGCTTTAATCTTTCCGAGATCTAGTTTTTCAG	3120
3121	Qy	CAAAAGCAGGATTTAAGAAATGTACTATCTTATGTGTGTTATGAAGAACATAGAAATCAAT	3180
3121	Db	CAAAAGCAGGATTTAAGAAATGTACTATCTTATGTGTGTTATGAAGAACATAGAAATCAAT	3180
3181	Qy	GCTGTATAAGTGCTTTTTTAACCTGTAAATTTTGTGAGGCTATCTTTATGCAATATAAT	3240
3181	Db	GCTGTATAAGTGCTTTTTTAACCTGTAAATTTTGTGAGGCTATCTTTATGCAATATAAT	3240
3241	Qy	ATTTTGAACATTTTACATTTGTTATATTTTTTAATCAGTTTTTACTCAAGTGATATATAC	3300
3241	Db	ATTTTGAACATTTTACATTTGTTATATTTTTTAATCAGTTTTTACTCAAGTGATATATAC	3300
3301	Qy	AAGAAAATGTAAACCACTGTATAGGGTAGAGTTATAAGAAATTTTGTCAAAATGTATTCACCCA	3360
3301	Db	AAGAAAATGTAAACCACTGTATAGGGTAGAGTTATAAGAAATTTTGTCAAAATGTATTCACCCA	3360
3361	Qy	TGTAGTCACCTCCTTTATGAAGAGACAGAACACGTCATCTCTCCAGAAAAGTTCCACAGTG	3420
3361	Db	TGTAGTCACCTCCTTTATGAAGAGACAGAACACGTCATCTCTCCAGAAAAGTTCCACAGTG	3420
3421	Qy	CTCCTTTTCCCTGAGTTTTCACAGTCTCTGGCAACAATGATCTGCTTCGTATTAATATAA	3480
3421	Db	CTCCTTTTCCCTGAGTTTTCACAGTCTCTGGCAACAATGATCTGCTTCGTATTAATATAA	3480
3481	Qy	CTGTTTCTAGATATTTGTAGCAATGTACCTTTTCCATATTTATTTTGTGTGTGTAAGGCTT	3540
3481	Db	CTGTTTCTAGATATTTGTAGCAATGTACCTTTTCCATATTTATTTTGTGTGTGTAAGGCTT	3540
3541	Qy	CTTTTGTAGTCATTAATAATTTTTCAGATTCATCTATGTTTAAATGTTCTTATCAGTAGTTGT	3600

[illegible]

RESULT 2

AL060804	185467 bp	linear	PRI 23-OCT-2002
LOCUS			
DEFINITION	Human DNA sequence from clone RP11-978115 on chromosome 1, complete sequence.		

ACCESSION AL606804

VERSION	AL60
1.0.0	1.0.0
1.0.1	1.0.1
1.0.2	1.0.2
1.0.3	1.0.3
1.0.4	1.0.4
1.0.5	1.0.5
1.0.6	1.0.6
1.0.7	1.0.7
1.0.8	1.0.8
1.0.9	1.0.9
1.0.10	1.0.10
1.0.11	1.0.11
1.0.12	1.0.12
1.0.13	1.0.13
1.0.14	1.0.14
1.0.15	1.0.15
1.0.16	1.0.16
1.0.17	1.0.17
1.0.18	1.0.18
1.0.19	1.0.19
1.0.20	1.0.20
1.0.21	1.0.21
1.0.22	1.0.22
1.0.23	1.0.23
1.0.24	1.0.24
1.0.25	1.0.25
1.0.26	1.0.26
1.0.27	1.0.27
1.0.28	1.0.28
1.0.29	1.0.29
1.0.30	1.0.30
1.0.31	1.0.31
1.0.32	1.0.32
1.0.33	1.0.33
1.0.34	1.0.34
1.0.35	1.0.35
1.0.36	1.0.36
1.0.37	1.0.37
1.0.38	1.0.38
1.0.39	1.0.39
1.0.40	1.0.40
1.0.41	1.0.41
1.0.42	1.0.42
1.0.43	1.0.43
1.0.44	1.0.44
1.0.45	1.0.45
1.0.46	1.0.46
1.0.47	1.0.47
1.0.48	1.0.48
1.0.49	1.0.49
1.0.50	1.0.50
1.0.51	1.0.51
1.0.52	1.0.52
1.0.53	1.0.53
1.0.54	1.0.54
1.0.55	1.0.55
1.0.56	1.0.56
1.0.57	1.0.57
1.0.58	1.0.58
1.0.59	1.0.59
1.0.60	1.0.60
1.0.61	1.0.61
1.0.62	1.0.62
1.0.63	1.0.63
1.0.64	1.0.64
1.0.65	1.0.65
1.0.66	1.0.66
1.0.67	1.0.67
1.0.68	1.0.68
1.0.69	1.0.69
1.0.70	1.0.70
1.0.71	1.0.71
1.0.72	1.0.72
1.0.73	1.0.73
1.0.74	1.0.74
1.0.75	1.0.75
1.0.76	1.0.76
1.0.77	1.0.77
1.0.78	1.0.78
1.0.79	1.0.79
1.0.80	1.0.80
1.0.81	1.0.81
1.0.82	1.0.82
1.0.83	1.0.83
1.0.84	1.0.84
1.0.85	1.0.85
1.0.86	1.0.86
1.0.87	1.0.87
1.0.88	1.0.88
1.0.89	1.0.89
1.0.90	1.0.90
1.0.91	1.0.91
1.0.92	1.0.92
1.0.93	1.0.93
1.0.94	1.0.94
1.0.95	1.0.95
1.0.96	1.0.96
1.0.97	1.0.97
1.0.98	1.0.98
1.0.99	1.0.99
1.0.100	1.0.100

KEYWORDS
HTG.
Homo sapiens (human)
SOURCE

SOURCE **HOMO SAPIENS (HUMAN)**
ORGANISM **Homo sapiens**

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primate
1 (1-200) 1 40 105467

REFERENCE AUTHORS

TITLE Direct Submission
JOURNAL Submitted (16-OCT-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: direct@wtsi.sanger.ac.uk

COMMENT

humquerysanger.ac.uk Clone requests: clonesrequest@sanger.ac.uk
On Oct 24, 2002 this sequence version replaced gi:15121511.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: En: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormpp This sequence was generated from part of bacterial clone contigs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at

http://www.sanger.ac.uk/HGP/Chr1
RP11-978115 is from the library RPC1-11.4 constructed by the group
of Pieter de Jong. For further details see
http://www.chori.org/hacnac/home.htm

VECTOR: pBACe3.6

----- Genome Center

Center: Wellcome Trust Sanger Ins
Center code: 60

Center code: SC
Web site: <http://www.sanger.ac>

11

FEATURES	
	Location/Qualifiers
1	105467

source
T. .18346/
/organism="Homo sapiens"

```

/mol_type="genomic DNA"

```

/db_xref="taxon:9606"

```
/chromosome="1"
/c1cnc="pp11-95"
```

/CLONE# RFI1-3/

Query Match	87.3%;	Score 3284;	DB 9;	Length 4032;
Best Local Similarity	99.9%;	Pred. No. 0;		
Matches 3604;	Conservative 0;	Mismatches 4;	Indels 1;	Gaps 1;
QY	113	GGCAGAAATGACTACATTTTGAAGAAAACTTCAAGATCAAGATAAGAAAGCAAGAAG	172	
Db	405	GGCAGAAATGACTACATTTTGAAGAAAACTTCAAGATCAAGATAAGAAAGCAAGAAG	464	
QY	173	TTTTCATCCACTTCTAATACGGAAAAACGAGAAATGGCAGTGGTTCTGAAGAAAGTGTGCTACA	232	
Db	465	TTTTCATCCACTTCTAATACGGAAAAACGAGAAATGGCAGTGGTTCTGAAGAAAGTGTGCTACA	524	
QY	233	CTGTCATTAATCACATCCCCCATCAGAGATCCCTCCCTGAGCTCCCAATGATGATGGCTATG	292	
Db	525	CTGTCATTAATCACATCCCCCATCAGAGATCCCTCCCTGAGCTCCCAATGATGATGGCTATG	584	
QY	293	AGAACATTTGACTCCCTCCACAGGAAAGTGAGACAGATTTAGAGAAAGGTTCAGACACAGAAT	352	
Db	585	AGAACATTTGACTCCCTCCACAGGAAAGTGAGACAGATTTAGAGAAAGGTTCAGACACAGAAT	644	
QY	353	ATGCGCTTCTTAGGACTTCTGTTAGTAGGCCCTTTGCTCTGCACCCATGAGCATGATATG	412	
Db	645	ATGCGCTTCTTAGGACTTCTGTTAGTAGGCCCTTTGCTCTGCACCCATGAGCATGATATG	704	
QY	413	AAGTTGTGTTTCACACTAAATTCCTCAAGCTGCTTTATCACCTTCACGCAATGAAGACA	472	
Db	705	AAGTTGTGTTTCACACTAAATTCCTCAAGCTGCTTTATCACCTTCACGCAATGAAGACA	764	
QY	473	ATGCAGAAATAGCAGACTCTCTGGGAGGTGTTACCTCGACGACGTGCAGTGAACATTCCTT	532	
Db	765	ATGCAGAAATAGCAGACTCTCTGGGAGGTGTTACCTCGACGACGTGCAGTGAACATTCCTT	824	
QY	533	CTCTGGCTTAAAGTTTGTAGAAATATATCTTATATATATCTTTAGGCAACTCTGATATGTGG	592	
Db	825	CTCTGGCTTAAAGTTTGTAGAAATATATCTTATATATATCTTTAGGCAACTCTGATATGTGG	884	
QY	593	CATCTCTGTGGCTTAGTGAAATCATAGAAATTGACA CAATGACCTTAAAAATATTTCTATGT	652	
Db	885	CATCTCTGTGGCTTAGTGAAATCATAGAAATTGACA CAATGACCTTAAAAATATTTCTATGT	944	
QY	653	GTATTTTCTGTGTAAGTTTTCAGGACATGGAGGTGATAAAAAACTTCTTAGGACAATA	712	
Db	945	GTATTTTCTGTGTAAGTTTTCAGGACATGGAGGTGATAAAAAACTTCTTAGGACAATA	1004	
QY	713	ATGTAAATGAAAATAAATTTCTAATCCCCCTGTACTAACTGAATGGACCTCTTTAGGC	772	
Db	1005	ATGTAAATGAAAATAAATTTCTAATCCCCCTGTACTAACTGAATGGACCTCTTTAGGC	1064	
QY	773	CAAAAGACCTCAGATGAACTGTGAAGACTGAAATCTGGCCATGATAGGAGGAGGTGA	832	
Db	1065	CAAAAGACCTCAGATGAACTGTGAAGACTGAAATCTGGCCATGATAGGAGGAGGTGA	1124	
QY	833	GACACACTTGTATTACCCCTTCCCTTTTGGAGTTTATGCAACAAGTGCACAGGATGAGTC	892	
Db	1125	GACACACTTGTATTACCCCTTCCCTTTTGGAGTTTATGCAACAAGTGCACAGGATGAGTC	1184	
QY	893	ATAAGACTGATGAATAGACTGATTTGGGCAATTAAGAGTCCCAATTCACACCTGACTCTG	952	
Db	1185	ATAAGACTGATGAATAGACTGATTTGGGCAATTAAGAGTCCCAATTCACACCTGACTCTG	1244	
QY	953	GTGTAGATCACACACTGTCTGAGGGATTCCCATCTATGAGACTTTTGTCTACATAACAGAGA	1012	
Db	1245	GTGTAGATCACACACTGTCTGAGGGATTCCCATCTATGAGACTTTTGTCTACATAACAGAGA	1304	
QY	1013	CTTTGGTTTCCAAACCCCTTTATTTTAGCTTAAAGCAATCTTTTCTACTGACTTCTTAAG	1072	
Db	1305	CTTTGGTTTCCAAACCCCTTTATTTTAGCTTAAAGCAATCTTTTCTACTGACTTCTTAAG	1364	
QY	1073	TCCTTTAGCAAGCTTAACTCTTCAACCAATTCGCAATCAGACAAACTTTTGAATCTTACC	1132	
Db	1365	TCCTTTAGCAAGCTTAACTCTTCAACCAATTCGCAATCAGACAAACTTTTGAATCTTACC	1424	
QY	1133	TATGACCTGTAAAGCTCTCTCCTGCTTCAAGATCTTGGCTCTTTAAGCTGAACCGATGTG	1192	

1425	Db	TATGACCTGTAAAGCTCTCTCCCTTGCTTCAAGATCTTGGCCCTCTTTAAGCTGACCGATGTGC	1484
1193	Qy	ACTTTCCATTTAATGANTATGCTCTTTGCTGTAACTCCTGTCTCCCTAAATATGTATAAA	1252
1485	Db	ACTTTCCATTTAATGANTATGCTCTTTGCTGTAACTCCTGTCTCCCTAAATATGTATAAA	1544
1253	Qy	AGTAAACGGTGAACCTGACCACTCAGGCACATTTTCTCAGGACCTCCTCAGAGTGTATCC	1312
1545	Db	AGTAAACGGTGAACCTGACCACTCAGGCACATTTTCTCAGGACCTCCTCAGAGTGTATCC	1604
1313	Qy	CAGGCCATGGTAAGTCAATGTGTGGCTCAGAATCAACCTCTTTAAATATTTTACAGAATTTG	1372
1605	Db	CAGGCCATGGTAAGTCAATGTGTGGCTCAGAATCAACCTCTTTAAATATTTTACAGAATTTG	1664
1373	Qy	GGTTTGGTTACCAATAAGTCTCCACAATAATATGTCCGAAGAACTTCAATTCGAAGCCT	1432
1665	Db	GGTTTGGTTACCAATAAGTCTCCACAATAATATGTCCGAAGAACTTCAATTCGAAGCCT	1724
1433	Qy	GCCTCACCAAAATTCAAATGCGCAACATCTCCCAATCCAAATTAACCTATTTCAATCTTTGAGGT	1492
1725	Db	GCCTCACCAAAATTCAAATGCGCAACATCTCCCAATCCAAATTAACCTATTTCAATCTTTGAGGT	1784
1493	Qy	GTAATCTACTCAATAAACTGTGTGAAGACGAGTGACAGACCCCTTTCGTACCTGACATTT	1552
1785	Db	GTAATCTACTCAATAAACTGTGTGAAGACGAGTGACAGACCCCTTTCGTACCTGACATTT	1844
1553	Qy	ACTTCAAAATTTCTTTTCTATGCTACTGGATATTTTTCGATATAAACTTCGACGTAATAGT	1612
1845	Db	ACTTCAAAATTTCTTTTCTATGCTACTGGATATTTTTCGATATAAACTTCGACGTAATAGT	1904
1613	Qy	TCAAAAATTAATAGTTTTTGACATTCGGCTTTCTGAGAAGAGAAATGAAAGTGTACAA	1672
1905	Db	TCAAAAATTAATAGTTTTTGACATTCGGCTTTCTGAGAAGAGAAATGAAAGTGTACAA	1964
1673	Qy	AATTAATAAAGATGAATGAAGCATATAAATGTCAATTTTTTCAATTTTCTAGTCAAC	1732
1965	Db	AATTAATAAAGATGAATGAAGCATATAAATGTCAATTTTTTCAATTTTCTAGTCAAC	2024
1733	Qy	AGAGAATCGAAGGATCTGTTCAAATATAGTAAAAAATGAAAAATAAACCTGTGCTTATA	1792
2025	Db	AGAGAATCGAAGGATCTGTTCAAATATAGTAAAAAATGAAAAATAAACCTGTGCTTATA	2084
1793	Qy	TTTTTGTTCGAAACACACTAGTTAATTTTAACTCTGCTAGTTATCTCTACCGAAGGTGA	1852
2085	Db	TTTTTGTTCGAAACACACTAGTTAATTTTAACTCTGCTAGTTATCTCTACCGAAGGTGA	2144
1853	Qy	TGTGTAGTTCTCGTTTAAAAATTCAGCAAACTGGAATAATCCATCTAAATATATGCTT	1912
2145	Db	TGTGTAGTTCTCGTTTAAAAATTCAGCAAACTGGAATAATCCATCTAAATATATGCTT	2204
1913	Qy	TCCTTCCAGAAGATTTTTTAATGATATGCGACGTTCTTAATTTGGAGACAAAGCCTTA	1972
2205	Db	TCCTTCCAGAAGATTTTTTAATGATATGCGACGTTCTTAATTTGGAGACAAAGCCTTA	2264
1973	Qy	ATTCACAATGCATTCATTATATATTTTTTGTATAGTTACAGTATACGAGTTCAGTATCC	2032
2265	Db	ATTCACAATGCATTCATTATATATTTTTTGTATAGTTACAGTATACGAGTTCAGTATCC	2324
2033	Qy	CTTAGATGAGATGCTTGGGACCGAAGTGTTTTGGATTTTCAGATTTATTTTTTGGATTTTG	2092
2325	Db	CTTAGATGAGATGCTTGGGACCGAAGTGTTTTGGATTTTCAGATTTATTTTTTGGATTTTG	2384
2093	Qy	GAATATTTCCATACATATAATGAGAGAGTTGGAAAAATGGGATTCAGCTCTAAATCATAAAA	2152
2385	Db	GAATATTTCCATACATATAATGAGAGAGTTGGAAAAATGGGATTCAGCTCTAAATCATAAAA	2444
2153	Qy	TTCACTTATGTTTGATATACACCTTATCTCGAATAGCCTGAAGGTAATTTTATACAAATTT	2212
2445	Db	TTCACTTATGTTTGATATACACCTTATCTCGAATAGCCTGAAGGTAATTTTATACAAATTT	2504
2213	Qy	TTAAATTAATTTTATGCTGTAAACAGAGTTTGGCCACATTCGACCATCAGAAAGCAGAGT	2272

Db 2505 TTAATAATATTTATGCTGGAACACAGAGTTTGGCCACATTGGACCATCAGAAAGCAGAAAGT 2564
QY 2273 GTCACTATTTCAAGTCAGTGTCAAAAAAGTTTCAGATGTTTAAGCTGGTATGCGATTCAT 2332
Db 2565 GTCACTATTTCAAGTCAGTGTCAAAAAAGTTTCAGATGTTTAAGCTGGTATGCGATTCAT 2624
QY 2333 GCCAGTGATCCAGTACTTTTGGGAAGCCAGACAGTGTGATCTCTTTGAGCCCGAGAGTTT 2392
Db 2625 GCGAGTGATCCAGTACTTTTGGGAAGCCAGACAGTGTGATCTCTTTGAGCCCGAGAGTTT 2684
QY 2393 GAGGCCAGACTGCACAACACAGTGTGAGACCTCGTTTCTACAATAATTAATAAAATTTAGCCA 2452
Db 2685 GAGGCCAGACTGCACAACACAGTGTGAGACCTCGTTTCTACAATAATTAATAAAATTTAGCCA 2744
QY 2453 GGTGTGGTGTGCACACCTGTAGTCCAGGTACTCAGGAGGCTGAGGTAGTAGGATGTT 2512
Db 2745 GGTGTGGTGTGCACACCTGTAGTCCAGGTACTCAGGAGGCTGAGGTAGTAGGATGTT 2804
QY 2513 TGAGACTGGGAGGTTGAGGCTGAACTGAGCCAGGATCTTGCCACACATTCAGAGCTTGGG 2572
Db 2805 TGAGACTGGGAGGTTGAGGCTGAACTGAGCCAGGATCTTGCCACACATTCAGAGCTTGGG 2864
QY 2573 CAACAGAGTGAACCTGTCTC-AAAAAAAAAAAAAGTTTCAGATTTTGGAGCATTTCA 2631
Db 2865 CAACAGAGTGAACCTGTCTC-AAAAAAAAAAAAAGTTTCAGATTTTGGAGCATTTCA 2924
QY 2632 GATCTTCAGATTAGGATTTTCAACCTGTACTGACCTTTTGTAGTCATTGACAGCATTAAT 2691
Db 2925 GATCTTCAGATTAGGATTTTCAACCTGTACTGACCTTTTGTAGTCATTGACAGCATTAAT 2984
QY 2692 CAATAGTGGAATCTGAGATAATCTATTTGCTGTATACACATTTGCTCTCTATTCAAG 2751
Db 2985 CAATAGTGGAATCTGAGATAATCTATTTGCTGTATACACATTTGCTCTCTATTCAAG 3044
QY 2752 AATTCATTATGCTCTGTGGTATTTAATGTGCGAAGGGAACAATAGAAATTTTGC 2811
Db 3045 AATTCATTATGCTCTGTGGTATTTAATGTGCGAAGGGAACAATAGAAATTTTGC 3104
QY 2812 AATTCAGAAAAGTCAATCTGTCAAAATATGTGCTGTAGTGTGATATGATATAAATAC 2871
Db 3105 AATTCAGAAAAGTCAATCTGTCAAAATATGTGCTGTAGTGTGATATGATATAAATAC 3164
QY 2872 GAAAAATGACAAAATTTTACTTTTCTGCTCTGCTTTGTAGTGTGATATGATATAAATAC 2931
Db 3165 GAAAAATGACAAAATTTTACTTTTCTGCTCTGCTTTGTAGTGTGATATGATATAAATAC 3224
QY 2932 CTATTATTGTAATAAAATTAATTTAATTTTGTAGTAACAATCTGGAATTTATCAGAGAAGGG 2991
Db 3225 CTATTATTGTAATAAAATTAATTTAATTTTGTAGTAACAATCTGGAATTTATCAGAGAAGGG 3284
QY 2992 CAGCAATAGGTTAATAAAGTATTTGATTTGTTAGGAGGAGGCTGGAATCCAGAGCAT 3051
Db 3285 CAGCAATAGGTTAATAAAGTATTTGATTTGTTAGGAGGAGGCTGGAATCCAGAGCAT 3344
QY 3052 CAATGCTCTTCTGCTGTGTTCCACATAAGCCACAGCAGATCTCTTAATCTTTCCGAGATCTTA 3111
Db 3345 CAATGCTCTTCTGCTGTGTTCCACATAAGCCACAGCAGATCTCTTAATCTTTCCGAGATCTTA 3404
QY 3112 GTTTTTCAGCAAGCAGGATTTAGAAATGTAACACTATCTATGTGTTATGAGAACAT 3171
Db 3405 GTTTTTCAGCAAGCAGGATTTAGAAATGTAACACTATCTATGTGTTATGAGAACAT 3464
QY 3172 AGAATCATGCTGTATAGTGTCTTTTAACTGTAAATTTTGTGAGCTTATCTTTTATG 3231
Db 3465 AGAATCATGCTGTATAGTGTCTTTTAACTGTAAATTTTGTGAGCTTATCTTTTATG 3524
QY 3232 CATATAAATTTTGAACATTTTACATGTTTATATTTTAAATCAGTTTACTCAAGTGTG 3291
Db 3525 CATATAAATTTTGAACATTTTACATGTTTATATTTTAAATCAGTTTACTCAAGTGTG 3584
QY 3292 ATTATATACAGAAAATGTAACCACTGTAGGAGTATGATTAAGAAATTTTGTCAAATGT 3351
Db 3585 ATTATATACAGAAAATGTAACCACTGTAGGAGTATGATTAAGAAATTTTGTCAAATGT 3644

QY 3352 ATTCAACCCATGTAGTCACTCTCTATGAAGAGACAGAACACATCATCTCTCCAGAAAGT 3411
Db 3645 ATTCAACCCATGTAGTCACTCTCTATGAAGAGACAGAACACATCATCTCTCCAGAAAGT 3704
QY 3412 TCCACAGTGTCTCTTTTCCCTCAGTTTCCACAGTCTCTGCAACCAACCAATGATCTGCTCGTA 3471
Db 3705 TCCACAGTGTCTCTTTTCCCTCAGTTTCCACAGTCTCTGCAACCAACCAATGATCTGCTCGTA 3764
QY 3472 TAATTAATACTGTCTAGATATTTGTAGCAATTTGTAGCAATTTGTAGCAATTTGTAGTGTG 3531
Db 3765 TAATTAATACTGTCTAGATATTTGTAGCAATTTGTAGCAATTTGTAGTGTG 3824
QY 3532 GTAAGGCTCTTTTGTAGTCAATTAATTTTGTAGATTCATCTATGTTTAAATGTTCTATC 3591
Db 3825 GTAAGGCTCTTTTGTAGTCAATTAATTTTGTAGATTCATCTATGTTTAAATGTTCTATC 3884
QY 3592 AGTAGTTGTACATCTTCTAGTCTCTAGCATATCCACATATAGATATATATTTGTTA 3651
Db 3885 AGTAGTTGTACATCTTCTAGTCTCTAGCATATCCACATATAGATATATATTTGTTA 3944
QY 3652 ATCTATCATCTGATGATGTAGGATATTTAGTATTTTGTAGATTTGATATTAAGTGGCT 3711
Db 3945 ATCTATCATCTGATGATGTAGGATATTTAGTATTTTGTAGATTTGATATTAAGTGGCT 4004
QY 3712 ATAAATGAA 3720
Db 4005 ATAAATGAA 4013

RESULT 4

AK124520 Homo sapiens cDNA FLJ42529 fis, clone BRACE3002390.
LOCUS AK124520
DEFINITION Homo sapiens
ACCESSION AK124520
VERSION 1 GI:34530324
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1
AUTHORS
Ninomiya, K., Wagatsuma, M., Kanda, K., Kondo, H., Yokoi, T.,
Kodaira, H., Furiya, T., Takahashi, M., Kikkawa, E., Omura, Y., Abe, K.,
Kaminara, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M.,
Sugiyama, T., Irie, K., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S.,
Yamanoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T.,
Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M.,
Kikuchi, H., Murakawa, K., Kanehori, K., Takahashi-Fujii, A.,
Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S.,
Nagahari, K., Maehuo, Y., Nagai, K. and Isogai, T.

NEDO human cDNA sequencing project

Unpublished

2 (bases 1 to 4136)

Isogai, T. and Yamamoto, J.

Direct Submission

Submitted (15-JUL-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7

Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan
(E-mail: genomics@hri.co.jp, Tel. 81-438-52-3975, Fax: 81-438-52-3986)

NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.

Location/Qualifiers

1. 4136

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="BRACE3002390"

CDS	ORIGIN	/tissue_type="cerebellum" /clone_lib="BRACE3" /note="Cloning vector: pME18SFL3" 136.537 /notes="unnamed protein product" /codon_start=1 /protein_id="BAC85873.1" /db_xref="GI:34530325" /translation="MWLSDSNSSCGGARNVCVCLVFLSVLGCQSERKGQMRTOQAG RMLRGASSETNPFGTLLIDPQSLPLAAGWGEHPVVRNVLPHAFSSWCPGQ NPFKXSETQELNRNVRTVKRWEIISSENSGRK"									
		Query Match	83.6%;	Score 3144;	DB 9;	Length 4136;					
		Best Local Similarity	99.8%;	Pred. No. 0;							
		Matches 3604;	Conservative	0;	Mismatches	2;	Indels	5;	Gaps	1;	
QY	113	GGCAGGAATGACTACATTTTGAAGAAGAAACTTCAAGATCAAGATATAGAAAGCCAGAAG	172								
DB	526	GGCAGGAATGACTACATTTTGAAGAAGAAACTTCAAGATCAAGATATAGAAAGCCAGAAG	585								
QY	173	TTTTCATCCACTTCTAATCAGAAAAACGAGAAATGGCAGTGGTCTGAAAGAAGTGTGCTACA	232								
DB	586	TTTTCATCCACTTCTAATCAGAAAAACGAGAAATGGCAGTGGTCTGAAAGAAGTGTGCTACA	645								
QY	233	CTGTCAATTAATCATATCCCATCAGAGATTCCTCCCTAGACTCCAATGATGATGGCTATG	292								
DB	646	CTGTCAATTAATCATATCCCATCAGAGATTCCTCCCTAGACTCCAATGATGATGGCTATG	705								
QY	293	AGAACATTGACTCCCTCACAAAGAAAGTGAAGAGTCTTATAGAGAAAGTCAAGACAGAAAT	352								
DB	706	AGAACATTGACTCCCTCACAAAGAAAGTGAAGAGTCTTATAGAGAAAGTCAAGACAGAAAT	765								
QY	353	ATGCCCTTCTTAGACTCTGTGATAGGCCTTTGTTCTCTGCACCCATGACCATGATATG	412								
DB	766	ATGCCCTTCTTAGACTCTGTGATAGGCCTTTGTTCTCTGCACCCATGACCATGATATG	825								
QY	413	AAAGTTGTGTTCCACACATAAAATCCTCAAGCTGCTTTATCACCTTCACAGCAATGAAGACA	472								
DB	826	AAAGTTGTGTTCCACACATAAAATCCTCAAGCTGCTTTATCACCTTCACAGCAATGAAGACA	885								
QY	473	ATCAGAAATAGCAGACTCTGGCGAAGTTGTCACCTGACGAGTGATGAACATTCCTT	532								
DB	886	ATCAGAAATAGCAGACTCTGGCGAAGTTGTCACCTGACGAGTGATGAACATTCCTT	945								
QY	533	TCTGGCTAAAGTTTAGAAAATATATCTATTATATATCTTCTTAGGCAACTCTGATATGTGG	592								
DB	946	TCTGGCTAAAGTTTAGAAAATATATCTATTATATATCTTCTTAGGCAACTCTGATATGTGG	1005								
QY	593	CATCTCTGTGGCTTAGTGAAATCATAGAAATGACACAAATGACCAATATTCCTATGT	652								
DB	1006	CATCTCTGTGGCTTAGTGAAATCATAGAAATGACCAATGACCAATATTCCTATGT	1065								
QY	653	GTTTTTCCTTTGTTAAAGTTTGAAGCATGGAGGTGATAAAAAACTTCTTAGGCAATA	712								
DB	1066	GTTTTTCCTTTGTTAAAGTTTGAAGCATGGAGGTGATAAAAAACTTCTTAGGCAATA	1125								
QY	713	ATGTAAATAGAAATAAATTTCTAATCCCTGACTAATGACCAATGACCTTCTTAGGC	772								
DB	1126	ATGTAAATAGAAATAAATTTCTAATCCCTGACTAATGACCAATGACCTTCTTAGGC	1185								
QY	773	CAAAAGACCTCAGATGAACCTGAAAGACTGAATTTCTGGCCATGATAGGAGGGAGTGA	832								
DB	1186	CAAAAGACCTCAGATGAACCTGAAAGACTGAATTTCTGGCCATGATAGGAGGGAGTGA	1245								
QY	833	GACACACTTGTATACCCCTTCCCTTTGGAGTTTATGCAACAGTGCACAGGATGATGC	892								
DB	1246	GACACACTTGTATACCCCTTCCCTTTGGAGTTTATGCAACAGTGCACAGGATGATGC	1305								
QY	893	ATAAGACTGATGAATAGACTGATTTGGCAATAAGAGTCCCAATTCACACTGACTCTG	952								
DB	1306	ATAAGACTGATGAATAGACTGATTTGGCAATAAGAGTCCCAATTCACACTGACTCTG	1365								

Db	2446	CTTAGATGAGTGTCTGGGACCAAGAGTGTCTGGATTTTCAGATTTTATTTTGGATTTTG	2505
Qy	2093	GAATATTTCCATACATATAATAGAGAGGTTGGAAATGGGATTCAGCTCAATCATATAAA	2152
Db	2506	GAATATTTCCATACATATAATAGAGAGGTTGGAAATGGGATTCAGCTCAATCATATAAA	2565
Qy	2153	TTCACTTATGTTTGAATACACCTTATCTGAATAGCCTGGAAGGTAATTTTATACAAATTT	2212
Db	2566	TTCACTTATGTTTGAATACACCTTATCTGAATAGCCTGGAAGGTAATTTTATACAAATTT	2625
Qy	2213	TTAAATAATTTTATGCCCTGAACACAGAGTTTGGGACATTTGGACCATCGAAAGCAGAAGT	2272
Db	2626	TTAAATAATTTTATGCCCTGAACACAGAGTTTGGGACATTTGGACCATCGAAAGCAGAAGT	2685
Qy	2273	GTCACTATTTCAAGTCAGTGTCTCAAAAAGTTTTCAGATGTTAAAGTGTGGTGCAGTTTCAT	2332
Db	2686	GTCACTATTTCAAGTCAGTGTCTCAAAAAGTTTTCAGATGTTAAAGTGTGGTGCAGTTTCAT	2745
Qy	2333	GGCAGTGATCCGAGTACTTTTGGGAAGCCAGACAGAGTGGATCTCTTGGAGCCAGGATTTT	2392
Db	2746	GGCAGTGATCCGAGTACTTTTGGGAAGCCAGACAGAGTGGATCTCTTGGAGCCAGGATTTT	2805
Qy	2393	GAGGCGAGCTGCACACACAGTGGAGACCTCGTTTCTCAAAATTAATTAATAAATTTAGCCA	2452
Db	2806	GAGGCGAGCTGCACACACAGTGGAGACCTCGTTTCTCAAAATTAATTAATAAATTTAGCCA	2865
Qy	2453	GGTGTGGTGGTGCACACCTGTAGTCCAGGTAAGTCTCAGAGGCTCAGGTAGGAGTGT	2512
Db	2866	GGTGTGGTGGTGCACACCTGTAGTCCAGGTAAGTCTCAGAGGCTCAGGTAGGAGTGT	2925
Qy	2513	TTGAGACTGGGAGTTGAGGCTGAACTGAGCCAGGATCTTGCACCAATTTCCAGCTTTGGG	2572
Db	2926	TTGAGACTGGGAGTTGAGGCTGAACTGAGCCAGGATCTTGCACCAATTTCCAGCTTTGGG	2985
Qy	2573	CAACAGAGTGAGACCTGTCTC- ----AAAAAAGTTTTCAGATTTTGGAGCAT	2627
Db	2986	CAACAGAGTGAGACCTGTCTC- ----AAAAAAGTTTTCAGATTTTGGAGCAT	3045
Qy	2628	TTCAATCTTCAATAGAGGATTTTCAACCTGTACTGACCTTTTGTAGTCAATGCAAGCAT	2687
Db	3046	TTCAATCTTCAATAGAGGATTTTCAACCTGTACTGACCTTTTGTAGTCAATGCAAGCAT	3105
Qy	2688	TAATCAATAGGTTGGACTCCAGATCACTCATTTTGTGTATACACATTTTGCCTCTCTATTC	2747
Db	3106	TAATCAATAGGTTGGACTCCAGATCACTCATTTTGTGTATACACATTTTGCCTCTCTATTC	3165
Qy	2748	AACGAATTTCTATGCCCTCTTGTGTGATTTTAAATGTCGGAAGGAAACAAATAGAAAT	2807
Db	3166	AACGAATTTCTATGCCCTCTTGTGTGATTTTAAATGTCGGAAGGAAACAAATAGAAAT	3225
Qy	2808	TTGCAATTTCTAGAAAGTCATTTCTGCAAAATATGTCAGTCTGTAGATATTAGCCAAAT	2867
Db	3226	TTGCAATTTCTAGAAAGTCATTTCTGCAAAATATGTCAGTCTGTAGATATTAGCCAAAT	3285
Qy	2868	TTAGAAAATGACAAAATTTTACTTTTCTGCTGCTTGTAGTCTTGTAGTCTTGTATGATATAA	2927
Db	3286	TTAGAAAATGACAAAATTTTACTTTTCTGCTGCTTGTAGTCTTGTAGTCTTGTATGATATAA	3345
Qy	2928	ATACCTTATTTGTATATAAATTTTAAATTTTGAATTAACCAATCTGGAATTTATCAGAGAA	2987
Db	3346	ATACCTTATTTGTATATAAATTTTAAATTTTGAATTAACCAATCTGGAATTTATCAGAGAA	3405
Qy	2988	GGGCAAGCAATAGGTTTAAATCAAGTATTTGATTTGGTGAAGGAAACGTTGAAATCCCAAGA	3047
Db	3406	GGGCAAGCAATAGGTTTAAATCAAGTATTTGATTTGGTGAAGGAAACGTTGAAATCCCAAGA	3465
Qy	3048	GCATCAATGCTTCTGTTGGTTCACCATAGCCACAGAGTGTCTTAATCTTTCCGAGA	3107
Db	3466	GCATCAATGCTTCTGTTGGTTCACCATAGCCACAGAGTGTCTTAATCTTTCCGAGA	3525
Qy	3108	TTCTAGTTTTTCAGCAAGCAGGATTTAAAGATGTAATCTTATGTTGTTTATGAAGAA	3167

Db	3526	TTCTAGTTTTTTCAGCAAGCAGGATTTAAGAAATGTAACATCTTATGTTGTTATGAAGAA	3585
Qy	3168	CAATAGATCATTTGCTGTATAGTCTCTTTTAACTGTAAATTTTGTGAAGCTTATCTTT	3227
Db	3586	CAATAGATCATTTGCTGTATAGTCTCTTTTAACTGTAAATTTTGTGAAGCTTATCTTT	3645
Qy	3228	TATGCATATAAATATTTTGAACATTTTACATTTGTTTATATTTTAACTAGTTTACTCAAG	3287
Db	3646	TATGCATATAAATATTTTGAACATTTTACATTTGTTTATATTTTAACTAGTTTACTCAAG	3705
Qy	3288	TGTGATTTATACAGAAATGTAACCACTCTTATGAAGAGACAGACAGTACATCTCCCAAGA	3347
Db	3706	TGTGATTTATACAGAAATGTAACCACTCTTATGAAGAGACAGACAGTACATCTCCCAAGA	3765
Qy	3348	ATGTATTTCAACCATGTAGTCACTCTCTTATGAAGAGACAGACAGTACATCTCCCAAGA	3407
Db	3766	ATGTATTTCAACCATGTAGTCACTCTCTTATGAAGAGACAGACAGTACATCTCCCAAGA	3825
Qy	3408	AGTTTCCACAGTCTCTTTTCCCTGAGTTTCCACAGTCTCTGGCAACCAATGATCTGTT	3467
Db	3826	AGTTTCCACAGTCTCTTTTCCCTGAGTTTCCACAGTCTCTGGCAACCAATGATCTGTT	3885
Qy	3468	CGTATATTTTAACTGTTTCTAGATATTTTGTAGCAATGTACCTTTCCATATTTTATTTGT	3527
Db	3886	CGTATATTTTAACTGTTTCTAGATATTTTGTAGCAATGTACCTTTCCATATTTTATTTGT	3945
Qy	3528	GTGTGTAAGGCTCTTTTGTAGTCAATTAATATTTTGTAGATTTTGTAGTCAATTAATTTT	3587
Db	3946	GTGTGTAAGGCTCTTTTGTAGTCAATTAATATTTTGTAGATTTTGTAGTCAATTAATTTT	4005
Qy	3588	TATCAGTGTGTGTACATCTTCTGTCAGCATATCACCATATAGATATATATATTTT	3647
Db	4006	TATCAGTGTGTGTACATCTTCTGTCAGCATATCACCATATAGATATATATATTTT	4065
Qy	3648	GTTAATCTAATCACTGATGATATGTAGATATTTTAAAGTTTGTAGATATTTTAAAGT	3707
Db	4066	GTTAATCTAATCACTGATGATATGTAGATATTTTAAAGTTTGTAGATATTTTAAAGT	4125
Qy	3708	GGCTATAATG 3718	
Db	4126	GGCTATAATG 4136	

RESULT 5

AC074365/c
LOCUS AC074365 141268 bp DNA linear HTG 23-SEP-2000
DEFINITION Homo sapiens chromosome 1 clone RP11-115C4, WORKING DRAFT SEQUENCE,
10 unordered pieces.
ACCESSION AC074365
VERSION AC074365.5 GI:10280935
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 141268)
2 (bases 1 to 141268)
The sequence of Homo sapiens clone
Unpublished
Waterston, R.H.
Direct Submission
Submitted (29-JUL-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Sep 23, 2000 this sequence version replaced gi:9838075.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc/index.shtml>
----- Project Information -----
Center project name: H_NH0115C04

----- Summary Statistics -----
Sequencing vector: M13; 100%
Chemistry: Dye-terminator; 100% of reads
Chemistry: Dye-terminator; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 136274 bases at least Q40
Consensus quality: 138356 bases at least Q30
Consensus quality: 139182 bases at least Q20
Insert size: 142000; agarose-fp
Insert size: 140368; sum-of-contigs
Quality coverage: 4.64 in Q20 bases; agarose-fp
Quality coverage: 4.81 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 10 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 31377: contig of 31377 bp in length
* 31378 31477: gap of unknown length
* 31478 31511: contig of 31511 bp in length
* 31511 31534: gap of unknown length
* 31534 31534: contig of 1734 bp in length
* 31534 31534: gap of unknown length
* 31534 31534: contig of 2661 bp in length
* 31534 31534: gap of unknown length
* 31534 31534: contig of 9968 bp in length
* 31534 31534: gap of unknown length
* 31534 31534: contig of 11482 bp in length
* 31534 31534: gap of unknown length
* 31534 31534: contig of 11044 bp in length
* 31534 31534: gap of unknown length
* 31534 31534: contig of 11415 bp in length
* 31534 31534: gap of unknown length
* 31534 31534: contig of 13698 bp in length
* 31534 31534: gap of unknown length
* 31534 31534: contig of 15455 bp in length.
* 125814 141268: contig of 15455 bp in length.

Location/Qualifiers
1. 141268
/organism="Homo sapiens"
/mol_type="Genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP11-115C4"
1. 31377
/note="assembly_name:Contig10"
31478. 63011
/note="assembly_name:Contig11"
63112. 64845
/note="assembly_name:Contig2"
64946. 67606
/note="assembly_name:Contig3"
67707. 77674
/note="assembly_name:Contig4"
77775. 89256
/note="assembly_name:Contig5"
89357. 100400
/note="assembly_name:Contig6"
100501. 111915
/note="assembly_name:Contig7"
112016. 125713
/note="assembly_name:Contig8"
125814. 141268
/note="assembly_name:Contig9"
clone_end:SP6
vector_side:left

FEATURES

source
1. 141268
/organism="Homo sapiens"
/mol_type="Genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP11-115C4"
1. 31377
/note="assembly_name:Contig10"
31478. 63011
/note="assembly_name:Contig11"
63112. 64845
/note="assembly_name:Contig2"
64946. 67606
/note="assembly_name:Contig3"
67707. 77674
/note="assembly_name:Contig4"
77775. 89256
/note="assembly_name:Contig5"
89357. 100400
/note="assembly_name:Contig6"
100501. 111915
/note="assembly_name:Contig7"
112016. 125713
/note="assembly_name:Contig8"
125814. 141268
/note="assembly_name:Contig9"
clone_end:SP6
vector_side:left

ORIGIN

Query Match 80.2%; Score 3016; DB 2; Length 141268;

Best Local Similarity 99.8%; Pred. No. 0;				Matches 3526; Conservative 0; Mismatches 3; Indels 3; Gaps 3;			
QY	190	CAGGAAACGAGAAATGCGAGTGGTCTTGAAGAAAGTGTGCTACCTGTCTATTAATCAATC	249				
DB	5700	CAGGAAACGAGAAATGCGAGTGGTCTTGAAGAAAGTGTGCTACCTGTCTATTAATCAATC	5641				
QY	250	CCCATCAGAGATCCTCCCTGAGCTCCAAATGATGCTGCTATGAGAAATGATGCTCCCTC	309				
DB	5640	CCCATCAGAGATCCTCCCTGAGCTCCAAATGATGCTGCTATGAGAAATGATGCTCCCTC	5581				
QY	310	ACAAGAAATGAGACAGATTTTGAAGAAAGTTCAGAGACAGAAATATGCCCTTCTTAGACT	369				
DB	5580	ACAAGAAATGAGACAGATTTTGAAGAAAGTTCAGAGACAGAAATATGCCCTTCTTAGACT	5521				
QY	370	TCTGTACTAGCCCTTCTGCTGACCCATGAGCATGATTAATGAAGTTGTGTTCCACAC	429				
DB	5520	TCTGTACTAGCCCTTCTGCTGACCCATGAGCATGATTAATGAAGTTGTGTTCCACAC	5461				
QY	430	TAAATCTCCTCAAGCTGCTTTATCACTCTTCCAGCAATGAAGCAATGAGCAATAGCAGCT	489				
DB	5460	TAAATCTCCTCAAGCTGCTTTATCACTCTTCCAGCAATGAAGCAATGAGCAATAGCAGCT	5401				
QY	490	CTGGGAAATGTTTCACTGAGCAGTGCATGAACATTCCTTCTGGCTAAAGTTAGA	549				
DB	5400	CTGGGAAATGTTTCACTGAGCAGTGCATGAACATTCCTTCTGGCTAAAGTTAGA	5341				
QY	550	AATATTAATCTTATATATATATATATCTTAGGCAATCTCTGATATGTGGCATCTCTCTGCTAGG	609				
DB	5340	AATATTAATCTTATATATATATATCTTAGGCAATCTCTGATATGTGGCATCTCTCTGCTAGG	5281				
QY	610	TGAATCATAGAAATGACAGATGACCTAAATATCTTATGTTGTTTGTGTTTAAAGT	669				
DB	5280	TGAATCATAGAAATGACAGATGACCTAAATATCTTATGTTGTTTGTGTTTAAAGT	5221				
QY	670	TTGAGGACATGGAGTGTATATAAAATCTTCTTAGGCAATATGTAATAATGAAATAA	729				
DB	5220	TTGAGGACATGGAGTGTATATAAAATCTTCTTAGGCAATATGTAATAATGAAATAA	5161				
QY	730	ATTCTAATCCCTGCTGATTAATGAGTGGACCTCTCTTAGGCCAAGAGACCTCAGATG	789				
DB	5160	ATTCTAATCCCTGCTGATTAATGAGTGGACCTCTCTTAGGCCAAGAGACCTCAGATG	5101				
QY	790	AACCTGAAAGACTGAATCTGCGCATGATGAGAGGAGGTGAGACACACCTTGTATAC	849				
DB	5100	AACCTGAAAGACTGAATCTGCGCATGATGAGAGGAGGTGAGACACACCTTGTATAC	5041				
QY	850	CCCTTCCCTTTGGAGTTTATGCAAGTACCAGGATGATGATGATGATGATGATGATGATG	909				
DB	5040	CCCTTCCCTTTTGGAGTTTATGCAAGTACCAGGATGATGATGATGATGATGATGATGATG	4981				
QY	910	GACTGATGTCGCAATAGAGTCCCAATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	969				
DB	4980	GACTGATGTCGCAATAGAGTCCCAATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	4921				
QY	970	TCTGAGGATTCATCTATGAGCTTTCTTACATAACAGAGACCTTGGTTTCCACAC	1029				
DB	4920	TCTGAGGATTCATCTATGAGCTTTCTTACATAACAGAGACCTTGGTTTCCACAC	4861				
QY	1030	CCTTTATTTAGCTAAGCTTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1089				
DB	4860	CCTTTATTTAGCTAAGCTTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	4801				
QY	1090	ACTCTTTCAACCAATGCGCAATC-AGACAACTTTTGAATCTTACCTATGACCTGTAAGCTC	1148				
DB	4800	ACTCTTTCAACCAATGCGCAATC-AGACAACTTTTGAATCTTACCTATGACCTGTAAGCTC	4742				
QY	1149	TCTCTGCTTCAAGATCTTGTGCTCTTTTAAAGTGAACCGATGCTGCTTCCATTTAATGA	1208				
DB	4741	TCTCTGCTTCAAGATCTTGTGCTCTTTTAAAGTGAACCGATGCTGCTTCCATTTAATGA	4682				
QY	1209	TTTATGCTCTTGTGTAACCT	1268				

Qy 3429 CCCTGAGTTTCACCAAGTCTGCTTGTATTAATAAAGTCTCTA 3488
 Db 2462 CCCTGAGTTTCACCAAGTCTGCTTGTATTAATAAAGTCTCTA 2403
 Qy 3489 GATATTTGTAGCAATGACCTTCCATATTTTGTGTGTAGGCTCTTTAGT 3548
 Db 2402 GATATTTGTAGCAATGACCTTCCATATTTTGTGTGTAGGCTCTTTAGT 2343
 Qy 3549 CATTATATATTTTGTAGATTCATCTATGTTTATGTTTCTATCATGATGTTGTACATCTTA 3608
 Db 2342 CATTATATATTTTGTAGATTCATCTATGTTTATGTTTCTATCATGATGTTGTACATCTTA 2283
 Qy 3609 CTTGTCTCAGCATATCACCATATAGATATATATAATTTGTTTAACTAATCATCTGATGGA 3668
 Db 2282 CTTGTCTCAGCATATCACCATATAGATATATATAATTTGTTTAACTAATCATCTGATGGA 2223
 Qy 3669 TATGTAGGATATTTAAGTTTGTGACATTTATGATTAATGATTAAGTGGCTATAATGAA 3720
 Db 2222 TATGTAGGATATTTAAGTTTGTGACATTTATGATTAATGATTAAGTGGCTATAATGAA 2171

RESULT 6
 BC024174 2648 bp mRNA linear PRI 04-OCT-2003
 LOCUS Homo sapiens LOC148823, mRNA (CDNA clone MGC:24564 IMAGE:4109064),
 DEFINITION complete cds.
 ACCESSION BC024174
 VERSION BC024174.1 Gi:18848218
 KEYWORDS MGC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 2648)
 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Haib, F.,
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.W., Hong, L.,
 Stapleton, K., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
 Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
 Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
 Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J.,
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
 Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
 Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y.,
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butterfield, Y.S., Kzyvinski, M.I., Skalska, U., Smalish, D.E.,
 Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16999-16903 (2002)
 22398257
 12477932
 2 (bases 1 to 2648)
 Strausberg, R.
 Direct Submission
 Submitted (19-FEB-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgaps-r@mail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: CLONTECH Laboratories, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
 contact: amadan@systemsbiology.org

Anup Madan, Jessica Fahey, Erin Helton, Mark Kettman, Anuradha
 Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 32 Row: B Column: 10
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 21687148.

Location/Qualifiers
 1..2648
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:24564 IMAGE:4109064"
 /tissue_type="Bone marrow, acute myelogenous leukemia"
 /clone_lib="NIH MGC_55"
 /lab_host="DH10B"
 /note="Vector: pDNR-LIB"
 1..2648
 /gene="LOC148823"
 /db_xref="LocusID:148823"
 48..455
 /codon_start=1
 /product="LOC148823"
 /protein_id="AAH24174.1"
 /db_xref="GI:18848218"
 /db_xref="LocusID:148823"
 /translation="MGNTLLRKLCLGENQKPKKGNDEERKROEMTTFERKLODOD
 KKSQVSTSTNQENSGSEVCTVINPHQRSSLSNDGYNIDSLTRKVRQF
 RERSSETVALLTSVSRPCSTHEHDYEVVPPH"

ORIGIN

Query Match 66.6%; Score 2506; DB 9; Length 2648;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 2606; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GAGAAACCGAGTCACGTGTGAAAGATGGAAATATATCTCTCGAAATATCATGTGCTG 60
 Db 24 GAGAAACCGAGTCACGTGTGAAAGATGGAAATATATCTCTCGAAATATCATGTGCTG 83
 Qy 61 GAGAGAAATCAAAAGAGCCCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 120
 Db 84 GAGAGAAATCAAAAGAGCCCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 143
 Qy 121 ATGACTACATTTTGAAGAGAAATCTTCAAGATCAAGATAAGAAAGCAAGAGTTTCATCC 180
 Db 144 ATGACTACATTTTGAAGAGAAATCTTCAAGATCAAGATAAGAAAGCAAGAGTTTCATCC 203
 Qy 181 ACTTCTAATCAGAAAGAAAGAGATGCGAGTGTCTTGAAGAGATGCTACACTGTCTATT 240
 Db 204 ACTTCTAATCAGAAAGAAAGAGATGCGAGTGTCTTGAAGAGATGCTACACTGTCTATT 263
 Qy 241 AATCACATCCCCATCAGAGATCTCTCCCTGAGCTCCCAATGATGATGCTATGAGAACATT 300
 Db 264 AATCACATCCCCATCAGAGATCTCTCCCTGAGCTCCCAATGATGATGCTATGAGAACATT 323
 Qy 301 GATCCCTCACAAGAGAAAGTGTAGAGATGTTAGAGAAAGTGTAGAGACAGAAATATGCCCC 360
 Db 324 GATCCCTCACAAGAGAAAGTGTAGAGATGTTAGAGAAAGTGTAGAGACAGAAATATGCCCC 383
 Qy 361 CTTAGGACTTCTCTTAGTAGGCTCTTCTCTGACCCATGAGCATGATTAATGAAGTTGTG 420
 Db 384 CTTAGGACTTCTCTTAGTAGGCTCTTCTCTGACCCATGAGCATGATTAATGAAGTTGTG 443
 Qy 421 TTTCCACACTAAATCTCTCAAGCTGTTTATACCTTCCAGCAATGAAGCAATGAGAA 480
 Db 444 TTTCCACACTAAATCTCTCAAGCTGTTTATACCTTCCAGCAATGAAGCAATGAGAA 503
 Qy 481 TAGAGACTCTGGGAGATGTTTACCCCTGAGCAGTGCATGAACATTTCTTCTGGCTA 540
 Db 504 TAGAGACTCTGGGAGATGTTTACCCCTGAGCAGTGCATGAACATTTCTTCTGGCTA 563
 Qy 541 AAGTTAGAAATATATCTTATTATATATCTTCTTAGGCAACTCTGATATGCGCATCTCTG 600

564	Db		AAAGTTTGAAGAATATTATCTTTATTATATATATCTTTAGGCAACTCTGATPATGTGGCATCTCTG	623
601	Qy		TGCGTTAGGTGAATCATACAAATTCACAAATGCACCAATGCCTAAATAATTTCTATGTGTTTTGCG	660
624	Db		TGGCTTAGGTGAATCATATAAATTCACCAATGCACCTAAATAATTTCTATGTGTTTTGCG	683
661	Qy		TTGTAAAGTTTGAAGCATGGAGTGATAAAAAAAATCTTCTTAGCAATAATGTAAAA	720
684	Db		TTGTAAAGTTTGAAGCATGGAGTGATAAAAAAAATCTTCTTAGCAATAATGTAAAA	743
721	Qy		TGAAATATAATTTCTAATCCCTGACTACTCAATGGACCTCTTCTAGGCCCAAGAGA	780
744	Db		TGAAATATAATTTCTAATCCCTGACTACTCAATGGACCTCTTCTAGGCCCAAGAGA	803
781	Qy		CCTCAGATGAACCTGAAAGACTGAATTTCTGGCCATGATAGGAAGGGAGGTGAGACACCC	840
804	Db		CCTCAGATGAACCTGAAAGACTGAATTTCTGGCCATGATAGGAAGGGAGGTGAGACACCC	863
841	Qy		TTGTTATACCGCTCCCTTTTGGAGTTTATGCAACAGTGACACAGGATGAGTCATAGACT	900
864	Db		TTGTTATACCGCTCCCTTTTGGAGTTTATGCAACAGTGACACAGGATGAGTCATAGACT	923
901	Qy		GATCAAAATAGACTGATTTGGCAATAGAGTCCCAATTCCAACTGACTCTGGGTAGAT	960
924	Db		GATCAAAATAGACTGATTTGGCAATAGAGTCCCAATTCCAACTGACTCTGGGTAGAT	983
961	Qy		CACACACTGTGAGGGAATCCATCTATGAGACATTTGTCTACATACAGAGACCTTGGTT	1020
984	Db		CACACACTGTGAGGGAATCCATCTATGAGACATTTGTCTACATACAGAGACCTTGGTT	1043
1021	Qy		TCCCAACCCCTTTATTTTAGCTTAAAGCATTTCTTCTACTGACTCTTTAAGTCTTTAGA	1080
1044	Db		TCCCAACCCCTTTATTTTAGCTTAAAGCATTTCTTCTACTGACTCTTTAAGTCTTTAGA	1103
1081	Qy		CAAAAGCTTAACCTTTTCAACCAATGGCAATFCAGACAAAATTTTGAATCTACCTTACCT	1140
1104	Db		CAAAAGCTTAACCTTTTCAACCAATGGCAATFCAGACAAAATTTTGAATCTACCTTACCT	1163
1141	Qy		GTAAGCTCTCTCGCTTCAGACTCTTTCAGCTCTTTCAGCTGAAACCGATGTGCACCTTCCA	1200
1164	Db		GTAAGCTCTCTCGCTTCAGACTCTTTCAGCTCTTTCAGCTGAAACCGATGTGCACCTTCCA	1223
1201	Qy		TTTAAATGATTTATGCTTTGCTTTGTAACCTCTCTCCCTCTCCCTAAAATGTTAAAAAGTAAACG	1260
1224	Db		TTTAAATGATTTATGCTTTGCTTTGTAACCTCTCTCCCTCTCCCTAAAATGTTAAAAAGTAAACG	1283
1261	Qy		GTGACCTGACCACTCTCAGGCACATCTTCTCAGGACCTCTGAGAGTGATCCCAGGCCAT	1320
1284	Db		GTGACCTGACCACTCTCAGGCACATCTTCTCAGGACCTCTGAGAGTGATCCCAGGCCAT	1343
1321	Qy		GGTAAGTCATGCTGGCTCAGAAATCAACCTCTTTAAATTAATTTTACAGAAATTTGGTTTTGG	1380
1344	Db		GGTAAGTCATGCTGGCTCAGAAATCAACCTCTTTAAATTAATTTTACAGAAATTTGGTTTTGG	1403
1381	Qy		TTACCAATAGTCTCCCAAAATATATGTCAGAGATCTTCAAATTCAGGCTCTCACCA	1440
1404	Db		TTACCAATAGTCTCCCAAAATATATGTCAGAGATCTTCAAATTCAGGCTCTCACCA	1463
1441	Qy		AAATTTCAAATGCAACATCTCCCATCAATTAATTAATTTTCACTATTTTCACTTTGAGGTTAATCTA	1500
1464	Db		AAATTTCAAATGCAACATCTCCCATCAATTAATTAATTTTCACTATTTTCACTTTGAGGTTAATCTA	1523
1501	Qy		CTCAATAAATCTGTGAAGACAGTGACAGACCCCTTTGCTTAACTGACATTTTACTTCAAT	1560
1524	Db		CTCAATAAATCTGTGAAGACAGTGACAGACCCCTTTGCTTAACTGACATTTTACTTCAAT	1583
1561	Qy		TTTTCTTTTCTCTATGCTACTGGATTTTTTGTGATATAAACTTTCAGTAAATAGTTCAAAAT	1620
1584	Db		TTTTCTTTTCTCTATGCTACTGGATTTTTTGTGATATAAACTTTCAGTAAATAGTTCAAAAT	1643
1621	Qy		TAATAGTTTTTGACATTCGGCTTTTCTGAGAGAGAGAAATTTGAAAGTGTCACAAAATAAAAA	1680

1644	TAATAGTTTTTGACATTTGGCTTTTCTGAGAGAGAAATTTGAAGTGTCTCAAAATATAAAA	1700
1681	AAGATGAAATGAGCATATATAATTTCTCAATTTTTTCAATTTTTTCTAGTCAACAGAGAAATC	1740
1704	AAGATGAAATGAGCATATATAATTTCTCAATTTTTTCAATTTTTTCTAGTCAACAGAGAAATC	1763
1741	GAGGATTCCTGTTCAATATTTAGTAAATATGGAATAAATCGAAATAAATCTGCTCTATATTTTGT	1800
1764	GAGGATTCCTGTTCAATATTTAGTAAATATGGAATAAATCTGCTCTATATTTTGT	1823
1801	GCAACACACTAGTTAAATTTAACTGTGACTAGTTATCTCTACGAAAGTGGATGTGTAGT	1860
1824	GCAACACACTAGTTAAATTTAACTGTGACTAGTTATCTCTACGAAAGTGGATGTGTAGT	1883
1861	TTCTGTTTTTAAATTTCAAGCAACTGGAAAAATATCCCATCTAAATATGCTTTCTTTCCC	1920
1884	TTCTGTTTTTAAATTTCAAGCAACTGGAAAAATATCCCATCTAAATATGCTTTCTTTCCC	1943
1921	AAGAAGTTTTTTAATGATATGCCAGCTTCCCTAATTTGGAGACAAAAGCCTTAATTGACAA	1980
1944	AAGAAGTTTTTTAATGATATGCCAGCTTCCCTAATTTGGAGACAAAAGCCTTAATTGACAA	2003
1981	TGCATTCATTTATATATTTTTTGTATAGTTACAGTATACAGTTAGTATCCCTTAGATG	2040
2004	TGCATTCATTTATATATTTTTTGTATAGTTACAGTATACAGTTAGTATCCCTTAGATG	2063
2041	AGATGCTTTGGGACACAGAAGTGTTTTGGATTTCCAGATTTATTTTTTGGATTTTGGATATTT	2100
2064	AGATGCTTTGGGACACAGAAGTGTTTTGGATTTCCAGATTTATTTTTTGGATTTTGGATATTT	2123
2101	CCATACATATATGAGAGAGTTGGAAAAATGGGATTTCAAGTCTAATCATAAAAATTCACCTTA	2160
2124	CCATACATATATATGAGAGAGTTGGAAAAATGGGATTTCAAGTCTAATCATAAAAATTCACCTTA	2183
2161	TGTTTGATATACACCTTATCTGAATAGCTTGAAGGTAAATTTTATACAATAATTTTAAATAA	2220
2184	TGTTTGATATACACCTTATCTGAATAGCTTGAAGGTAAATTTTATACAATAATTTTAAATAA	2243
2221	TTTTATGCCTGAAACAGAGTTTGGCACAATTTGAACCATCAGAAAGCAGAGTGTCACTAT	2280
2244	TTTTATGCCTGAAACAGAGTTTGGCACAATTTGAACCATCAGAAAGCAGAGTGTCACTAT	2303
2281	TTCAAGTCAGTGTCTCAAAAAGTTTCAGATCTTAAGCTTGGTGATGCGAGTTCAATGCCAGTGA	2340
2304	TTCAAGTCAGTGTCTCAAAAAGTTTCAGATCTTAAGCTTGGTGATGCGAGTTCAATGCCAGTGA	2363
2341	TCCGAGTACTTTGGGAAACCAAGACAGAGTGGATCTCTTTGAGCCGACAGAGTTTGAGGCCAG	2400
2364	TCCGAGTACTTTGGGAAACCAAGACAGAGTGGATCTCTTTGAGCCGACAGAGTTTGAGGCCAG	2423
2401	ACTGCACACACAGTGTAGACCTCTGTTTCTACAAATAATTAATAAATTTAGCCAGTGTGGT	2460
2424	ACTGCACACACAGTGTAGACCTCTGTTTCTACAAATAATTAATAAATTTAGCCAGTGTGGT	2483
2461	GGTGGACACCTGTATGCTCCAGGTACTCAGAGAGCTGAGGTAGTAGGATTTGTGTGAGACTG	2520
2484	GGTGGACACCTGTATGCTCCAGGTACTCAGAGAGCTGAGGTAGTAGGATTTGTGTGAGACTG	2543
2521	GGAGTTGAGGCTGAACTGAGCCACAGGATCTTGCCACACATTCACAGCTTGGGCACACAGAG	2580
2544	GGAGTTGAGGCTGAACTGAGCCACAGGATCTTGCCACACATTCACAGCTTGGGCACACAGAG	2603
2581	TGAGACCTCTCTCAAAAAAATAAAAAA	2608
2604	TGAGACCTCTCTCAAAAAAATAAAAAA	2631

RESULT 7
AK123798
LOCUS
DEFINITION
ACCESSION
VERSION

KEYWORDS

oligo capping; fis (full insert sequence).

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

AUTHORS

Suzuki, O., Sasaki, N., Aotsuka, S., Shoji, T., Ichihara, T., Shiohata, N., Matsumoto, K., Hirano, M., Sano, S., Nomura, R., Yoshikawa, Y., Matsumura, Y., Moriya, S., Chiba, E., Momiya, H., Onogawa, S., Kaeriyama, S., Satoh, N., Matsunawa, H., Takahashi, E., Kataoka, R., Kuga, N., Kuroda, A., Satoh, I., Kamata, K., Takami, S., Teraehima, Y., Watanabe, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Matsui, H., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Wagatsuma, M., Murakawa, K., Kanehori, K., Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahari, K., Masuo, Y., Nagai, K. and Isogai, T.

TITLE

NEDO human cDNA sequencing project

JOURNAL

Unpublished

REFERENCE

AUTHORS

Isogai, T. and Yamamoto, J.

TITLE

Direct Submission

JOURNAL

Submitted (15-JUL-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986) NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5' & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.

FEATURES

source

1. 1709
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clones="NOVAR2000710"
/tissue_type="ovary"
/clone_lib="NOVAR2"
/dev_stage="adult"
/note="cloning vector: pME18SFL3"

ORIGIN

Query Match 43.2%; Score 1624; DB 9; Length 1709;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1674; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGAAACCGAGTCACTGTGAAAAGATGGGAAATATCTCTCGGAAAACTCAGTTGGCTG 60
DB 35 GAGAAACCGAGTCACTGTGAAAAGATGGGAAATATCTCTCGGAAAACTCAGTTGGCTG 94
QY 61 GGAGAGATCAAAAGAACCCAGAAAGAACCCAGATGAGGAAGAAAGAAACCGCAGGAA 120
DB 95 GGAGAGATCAAAAGAACCCAGAAAGAACCCAGATGAGGAAGAAAGAAACCGCAGGAA 154
QY 121 ATGACTCATTTGAAAGAAAACCTCAAGATCAAGATAAGAAAAGCCAGAAAGTTTCATCC 180
DB 155 ATGACTCATTTGAAAGAAAACCTCAAGATCAAGATAAGAAAAGCCAGAAAGTTTCATCC 214
QY 181 ACTTCTAATCAGGAAAACGAGAAATGGCAGTGGTTCTGAAGAAGTGTGTACACTGTCAAT 240
DB 215 ACTTCTAATCAGGAAAACGAGAAATGGCAGTGGTTCTGAAGAAGTGTGTACACTGTCAAT 274
QY 241 AATCATATCCCCCATCAGAGATCCCTCCCTGAGCTCCATGATGATGGCTATGAGAACATT 300
DB 275 AATCATATCCCCCATCAGAGATCCCTCCCTGAGCTCCATGATGATGGCTATGAGAACATT 334
QY 301 GACTCCCTCACAAGGAAAGTGAAGACAGTTTAGAGAAAAGGTGAGAGACAGAAATATGCCCTT 360
DB 335 GACTCCCTCACAAGGAAAGTGAAGACAGTTTAGAGAAAAGGTGAGAGACAGAAATATGCCCTT 394

QY 361 CTTAGGACTTCTGTGTAGTAGGCCTTGTTCCTGCACCCCATGAGCATGATTTAAGATTGTG 420
DB 395 CTTAGGACTTCTGTGTAGTAGGCCTTGTTCCTGCACCCCATGAGCATGATTTAAGATTGTG 454
QY 421 TTTCACACATAAATCTCAAGCTGCTTTATCATCCTCCAGCAATGAGAGCAATGCAGAA 480
DB 455 TTTCACACATAAATCTCAAGCTGCTTTATCACCCTCCAGCAATGAGAGCAATGCAGAA 514
QY 481 TAGCAGACTCTGGCGAAGTTGTTCACCTCGAGCAGTGCATGAAACATTCTCTTTCTGGCTA 540
DB 515 TAGCAGACTCTGGCGAAGTTGTTCACCTCGAGCAGTGCATGAAACATTCTCTTTCTGGCTA 574
QY 541 AAGTTTAGAATATTATCTTATTATATATCTTAGGCAACTCTGATATGTGGCATCTCTG 600
DB 575 AAGTTTAGAATATTATCTTATTATATATCTTAGGCAACTCTGATATGTGGCATCTCTG 634
QY 601 TGGCTTAGGTGAAATCATAGAAATGACACAATGACCTTAAATAATTTCTATGTGTTTTGC 660
DB 635 TGGCTTAGGTGAAATCATAGAAATGACACAATGACCTTAAATAATTTCTATGTGTTTTGC 694
QY 661 TTGTAAGTTTGAAGACATGGAGGTGATAAAAAAATTTCTTAGGCAATAATGTAATAA 720
DB 695 TTGTAAGTTTGAAGACATGGAGGTGATAAAAAAATTTCTTAGGCAATAATGTAATAA 754
QY 721 TGAAAATAAATTTCTAAATCCCTGACTAACTGAATGACCTCTCTTAGGCCAAAGAGA 780
DB 755 TGAAAATAAATTTCTAAATCCCTGACTAACTGAATGACCTCTCTTAGGCCAAAGAGA 814
QY 781 CCTCAGATGAACCTGAAAGACTGAATTTCTGGCCATGATAGGAAGGAGGTGAGACACACC 840
DB 815 CCTCAGATGAACCTGAAAGACTGAATTTCTGGCCATGATAGGAAGGAGGTGAGACACACC 874
QY 841 TTGTTATACCCCTTCCCTTTTGGAGTTTATGCACAAGTGACCAGGATGAGTCATAGACT 900
DB 875 TTGTTATACCCCTTCCCTTTTGGAGTTTATGCACAAGTGACCAGGATGAGTCATAGACT 934
QY 901 GATGAATAGACTGATTTGGCAATAAGAGTCCCAATCCAACTGACTCTGGTGTAGAT 960
DB 935 GATGAATAGACTGATTTGGCAATAAGAGTCCCAATCCAACTGACTCTGGTGTAGAT 994
QY 961 CACACACTGTCTGAGGATTCCTATGAGACTTTGTCTACATACAGAGACCTTGGTT 1020
DB 995 CACACACTGTCTGAGGATTCCTATGAGACTTTGTCTACATACAGAGACCTTGGTT 1054
QY 1021 TCCACAACCCCTTTATTTTAGCTAAAGCAATCTTTTCTACTGACTCTTTAAGTCTTTAGA 1080
DB 1055 TCCACAACCCCTTTATTTTAGCTAAAGCAATCTTTTCTACTGACTCTTTAAGTCTTTAGA 1114
QY 1081 CAAAGCTTAACCTTTCAACCAATTCGCAATCAGACAAACTTTGAATCTACCTATGACCT 1140
DB 1115 CAAAGCTTAACCTTTCAACCAATTCGCAATCAGACAAACTTTGAATCTACCTATGACCT 1174
QY 1141 GTAAGCTCTCTCTGCTCAAGATCTTGCTCTTTAAGCTGAACCGATGTCACCTTTCCA 1200
DB 1175 GTAAGCTCTCTCTGCTCAAGATCTTGCTCTTTAAGCTGAACCGATGTCACCTTTCCA 1234
QY 1201 TTTAATGATTTATGTTTGTGTGTAACTCTCTGTCTCCCTAAATGTTATAAAGTAAAGC 1260
DB 1235 TTTAATGATTTATGTTTGTGTGTAACTCTCTGTCTCCCTAAATGTTATAAAGTAAAGC 1294
QY 1261 GTGACCTGACCACTCAGGACACACTTTCTCAGGACCTCCTGAGAGTGTATCCAGGCCAT 1320
DB 1295 GTGACCTGACCACTCAGGACACACTTTCTCAGGACCTCCTGAGAGTGTATCCAGGCCAT 1354
QY 1321 GGTAACTCATGTGGCTCAGAAATCAACCTCTTTAAATATTTTACAGAAATTTGGTTTTGG 1380
DB 1355 GGTAACTCATGTGGCTCAGAAATCAACCTCTTTAAATATTTTACAGAAATTTGGTTTTGG 1414
QY 1381 TTACCAATAGTCTCCACAATAATATGTCCAGAAATCTTCAATTTCCAAGCTGCTCACCAC 1440
DB 1415 TTACCAATAGTCTCCACAATAATATGTCCAGAAATCTTCAATTTCCAAGCTGCTCACCAC 1474

```

Qy 1441 AATTTCAAATGCCAATCTCCCATCAATTAATCTATTTTCATCTTTGAGGTGTAATCTA 1500
Db 1475 AATTTCAAATGCCAATCTCCCATCAATTAATCTATTTTCATCTTTGAGGTGTAATCTA 1534
Qy 1501 CTCATTAATCTGTGTAGACAGAGTGACAGACCCCTTGGCTAACTGACATTTACTTCAAT 1560
Db 1535 CTCATTAATCTGTGTAGACAGAGTGACAGACCCCTTGGCTAACTGACATTTACTTCAAT 1594
Qy 1561 TTTTCCTTTTCTATGTACTGATATTTTTCATATAAACTTGCAGTAACTAGTTCAAAAAT 1620
Db 1595 TTTTCCTTTTCTATGTACTGATATTTTTCATATAAACTTGCAGTAACTAGTTCAAAAAT 1654
Qy 1621 TAAATAGTTTTCGATCGCTTTTCTGAGAGAGAGAAATGAAAGTGTCACAAAAT 1675
Db 1655 TAAATAGTTTTCGATCGCTTTTCTGAGAGAGAGAAATGAAAGTGTCACAAAAT 1709

RESULT 8
AC023386
LOCUS Homo sapiens chromosome 1 clone RP11-482N10 map 1, LOW-PASS
DEFINITION SEQUENCE SAMPLING.
ACCESSION AC023386
VERSION AC023386.2 GI:74117784
KEYWORDS HTG; HTGS PHASE0.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 121353)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 1, clone RP11-482N10
Unpublished
2 (bases 1 to 121353)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Beda,F., Boguslavskiy,L.,
Boukhgalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,
Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
Dearellano,K., Dewar,K., Dodge,S., Domino,M., Doyle,M.,
Penestor,J., Ferreira,P., FitzHugh,W., Forrest,C., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., Landers,T., Largocque,K., Lehoczy,J., Levine,R.,
Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N., McCarthy,M.,
McSwan,P., McGurk,A., McKernan,K., McPheeters,R., Meldrim,J.,
Meneus,L., Mihova,T., Miranda,C., Mienga,V., Morrow,J., Naylor,J.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,T.M.,
Peterson,K., Pierre,N., Pisanic,C., Pollara,V., Raymond,C.,
Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S.,
Severy,P., Spencer,B., Stange-Thomann,N., Tesfaye,S., Tirrell,A.,
Subramanian,A., Talamas,J., Tefaye,S., Theodore,J., Tirrell,A.,
Traversa,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B.,
Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and
Zody,M.
Direct Submission
Submitted (14-FEB-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 121353)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Boguslavskiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collymore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
Dodgson,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., Larocque,K., Lamazares,R., Landers,T., Lehoczy,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
McCarthy,M., McSwan,P., McGurk,A., McKernan,K., McPheeters,R.,
Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mienga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,

```

```

O'Neil,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tesfaye,S., Theodore,J., Tirrell,A., Traversa,M., Trigilio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 5, 2000 this sequence version replaced gi:6970536.
All repeats were identified using RepeatMasker:
Smit,A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L5247
Center clone name: 482_N_10
-----

```

```

* NOTE: This record contains 134 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

```

```

* 1 798: contig of 798 bp in length
* 799 898: gap of 100 bp
* 899 1681: contig of 783 bp in length
* 1682 1781: gap of 100 bp
* 1782 2593: contig of 812 bp in length
* 2594 2894: gap of 100 bp
* 2894 3516: contig of 823 bp in length
* 3517 3617: gap of 100 bp
* 3617 4437: contig of 820 bp in length
* 4437 5364: gap of 100 bp
* 5365 5464: contig of 828 bp in length
* 5465 6260: contig of 796 bp in length
* 6261 6360: gap of 100 bp
* 6361 7174: contig of 813 bp in length
* 7174 7273: gap of 100 bp
* 7274 8067: contig of 794 bp in length
* 8068 8167: gap of 100 bp
* 8168 8958: contig of 791 bp in length
* 8959 9059: gap of 100 bp
* 9059 9867: contig of 809 bp in length
* 9868 9968: gap of 100 bp
* 9968 10764: contig of 797 bp in length
* 10765 11687: contig of 823 bp in length
* 11688 11787: gap of 100 bp
* 11788 12616: contig of 829 bp in length
* 12617 13336: contig of 820 bp in length
* 13337 14559: gap of 100 bp
* 14560 14559: contig of 823 bp in length
* 14560 15356: contig of 797 bp in length
* 15357 15456: gap of 100 bp
* 15457 16247: contig of 791 bp in length
* 16248 17339: gap of 100 bp
* 17339 17239: contig of 792 bp in length
* 17140 17239: gap of 100 bp
* 17240 18022: contig of 783 bp in length

```

```

TITLE
JOURNAL
COMMENT

```


Boukhgalter, B., Brown, A., Burkett, G., Campopiano, A., Castle, A.,
Chesepel, Y., Colangelo, M., Collins, S., Collimore, A., Cooke, P.,
Dearellano, K., Dewar, K., Dodge, S., Dominio, M., Doyle, M.,
Fenstermaker, J., Ferreira, P., Fitzhugh, M., Forrest, C., Gage, D.,
Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
Grand-Pierre, N., Grant, G., Hagsos, B., Heaford, A., Horton, L.,
Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,
Klein, J., Landers, T., Largocque, K., Lehoczyk, J., Levine, R.,
Lieu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N., McCarthy, M.,
McEwan, P., McGurk, A., McKernan, K., McPheeters, R., Meldrim, J.,
Meneus, P., Mihova, T., Miranda, C., Mlenga, V., Morrow, J., Naylor, J.,
Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Olivari, T.M.,
Peterson, K., Pierre, N., Pisan, C., Pollara, V., Raymond, C.,
Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S.,
Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Tirrell, A.,
Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B.,
Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zimmer, A. and
Zody, M.

Direct Submission
Submitted (14-FEB-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 121353)
Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F.,
Boguslavsky, L., Boukhgalter, B., Brown, A., Burkett, G.,
Campopiano, A., Castle, A., Chesepel, Y., Colangelo, M., Collins, S.,
Collimore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S.,
Dodge, S., Dominio, M., Doyle, M., Ferreira, P., Fitzhugh, M., Gage, D.,
Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
Grand-Pierre, N., Grant, G., Hagsos, B., Heaford, A., Horton, L.,
Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,
Klein, J., Largocque, K., Lamazares, R., Landers, T., Lehoczyk, J.,
Levine, R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N.,
McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R.,
Meldrim, J., Meneus, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J.,
Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
O'Neill, D., Olivari, T.M., Oliver, J., Peterson, K., Pierre, N.,
Pisan, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Tesfaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J.,
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,
Young, G., Zainoun, J., Zimmer, A. and Zody, M.

Direct Submission
Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 5, 2000 this sequence version replaced gi:5970536.
All repeats were identified using RepeatMasker:
Smit, A.P.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L5247
Center clone name: 482_N_10

NOTE: This record contains 134 individual
sequencing reads that have not been assembled into
contigs. Runs of N are used to separate the reads
and the order in which they appear is completely
arbitrary. Low-pass sequence sampling is useful for
identifying clones that may be gene-rich and allows
overlap relationships among clones to be deduced.
However, it should not be assumed that this clone
will be sequenced to completion. In the event that
the record is updated, the accession number will
be preserved.

798: contig of 798 bp in length
898: gap of 100 bp

1681: contig of 783 bp in length
1781: gap of 100 bp
2593: contig of 812 bp in length
2594: gap of 100 bp
3516: contig of 823 bp in length
3517: gap of 100 bp
4436: contig of 820 bp in length
4536: gap of 100 bp
5364: contig of 828 bp in length
5464: gap of 100 bp
6260: contig of 796 bp in length
6360: gap of 100 bp
7173: contig of 813 bp in length
7273: gap of 100 bp
8067: contig of 794 bp in length
8167: gap of 100 bp
8958: contig of 791 bp in length
9058: gap of 100 bp
9667: contig of 809 bp in length
9667: gap of 100 bp
10764: contig of 797 bp in length
10864: gap of 100 bp
11687: contig of 823 bp in length
11787: gap of 100 bp
12616: contig of 829 bp in length
12716: gap of 100 bp
13536: contig of 820 bp in length
13636: gap of 100 bp
14459: contig of 823 bp in length
14559: gap of 100 bp
15356: contig of 797 bp in length
15456: gap of 100 bp
16247: contig of 791 bp in length
16347: gap of 100 bp
17139: contig of 792 bp in length
17239: gap of 100 bp
18022: contig of 783 bp in length
18122: gap of 100 bp
18924: contig of 802 bp in length
19024: gap of 100 bp
19823: contig of 805 bp in length
19923: gap of 100 bp
20723: contig of 794 bp in length
20823: gap of 100 bp
21619: contig of 796 bp in length
21719: gap of 100 bp
22531: contig of 812 bp in length
22631: gap of 100 bp
23459: contig of 828 bp in length
23559: gap of 100 bp
24373: contig of 814 bp in length
24473: gap of 100 bp
25288: contig of 815 bp in length
25388: gap of 100 bp
26132: contig of 804 bp in length
26232: gap of 100 bp
27089: contig of 797 bp in length
27189: gap of 100 bp
28007: contig of 818 bp in length
28107: gap of 100 bp
28918: contig of 811 bp in length
29018: gap of 100 bp
29820: contig of 802 bp in length
29920: gap of 100 bp
30737: contig of 817 bp in length
30837: gap of 100 bp
31670: contig of 833 bp in length
31770: gap of 100 bp
32585: contig of 815 bp in length
32685: gap of 100 bp
33496: contig of 811 bp in length
33596: gap of 100 bp
34397: contig of 801 bp in length

899
1682
1782
2594
2594
3517
3517
4437
4537
5365
5465
6261
6361
7174
7274
8068
8168
8959
9059
9668
10765
11688
11788
12617
12717
13537
13637
14460
14560
15357
15457
16248
16348
17140
17240
18023
18123
18925
19025
19830
19930
20724
20824
21620
21720
22532
22632
23459
23559
24374
24474
25289
25389
26133
26233
27090
27190
28008
28108
28919
29019
29821
29921
30738
30838
31671
31771
32586
32686
33497
33597

Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (31-JAN-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 144333)
Birnbaum,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boquelavkiy,L., Boukagater,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B., Dearellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Fato,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Harez,Z.N., Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., Maclean,C., MacDonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrim,J., Menes,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schuback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliou,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission
Submitted (13-MAR-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 13, 2003 this sequence version replaced gi|27753682.
All repeats were identified using RepeatMasker:
Smit, A.P.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L23019
Center clone name: 2323_K_18

FEATURES
Source
1..144333
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="15"
/map="15"
/clone="CTD-2323K18"
/clone_lib="CITDI Human BAC"
543..640
/rpt_family="WIR"
647..921
/rpt_family="AluSc"
complement(1163..1273)
/rpt_family="AluY"
complement(1442..1741)
/rpt_family="AluSp"
complement(1823..1962)
/rpt_family="FLAM_C"
2221..2407
/rpt_family="L2"
2435..2462
/rpt_family="AT_rich"
3336..3478
/rpt_family="AluJo"
3490..3594
/rpt_family="AluJo"
complement(3711..3754)
/rpt_family="MIR3"
complement(4266..4574)
/rpt_family="AluSq"
5366..5468

AUTHORS Klein, S.L., Strausberg, R.L., Wagner, L., Pontius, J., Clifton, S.W. and Richardson, P.

TITLE Genetic and genomic tools for Xenopus research: The NIH Xenopus Initiative

JOURNAL MEDLINE Dev. Dyn. 225 (4), 384-391 (2002)

PUBMED 22341132

REFERENCE 12454917

AUTHORS 2 (bases 1 to 784)

Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Ustin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalls, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.

TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

JOURNAL MEDLINE Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

PUBMED 22388257

REFERENCE 12477932

AUTHORS 3 (bases 1 to 784)

Klein, S. and Strausberg, R.

TITLE Direct Submission

JOURNAL Submitted (01-MAY-2003) National Institutes of Health, Xenopus Gene Collection (XGC), National Institute of Child Health and Human Development, 6100 Executive Boulevard, Room 4B01, Rockville, MD 20892-7510, USA

REMARK COMMENT NIH-MGC Project

Contact: XGC help desk

Email: cgapbs@mail.nih.gov

Tissue Procurement: Dr. Igor Dawid

cDNA Library Prepared by: The I.M.A.G.E. Consortium, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada

info@cgsc.bc.ca

Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Scott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

FEATURES source

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAK plate: 95 Row: k Column: 3

This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein.

Location/Qualifiers

1. .784

/db_xref="taxon:8355"

/mol_type="mRNA"

/organism="Xenopus laevis"

/clone="MGC:53762 IMAGE:5537106"

/tissue_type="embryo, stage 31/32, Xenopus"

/cclone_lib="NICHG_XGC_Emb4"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

1. .784

gene

CDS /gene="MGC53762"

/db_xref="LocusID:379407"

19..513

/codon_start=1

/product="MGC53762 protein"

/protein_id="AAH52102.1"

/db_xref="GI:30704995"

/db_xref="LocusID:379407"

/translation="MESVLSFVLLVATSTFLAKEIKPKVSKPOTLSRGWGDNLWV QYVRSGLYKAKAENKPLMLNHNDCPHSLAKKAFASHQGKLAEPILLNVYDP TKNQQLDQGVYFKLIIFVDFSLVVRADLPFGKISNHYTFADILHLENKKALVLL KTEL"

ORIGIN

Query Match 1.4%; Score 52; DB 5; Length 784;

Best Local Similarity 100.0%; Pred. No. 1.9e-14;

Matches 52; Conservative 0; Mismatches 0; Gaps 0;

QY 3711 TATAATGAAA 3762

Db 718 TATAATGAAA 769

RESULT 14

BC009760 1815 bp mRNA linear PRI 30-SEP-2003

LOCUS Homo sapiens chromosome 4 open reading frame 9, mRNA (CDNA clone IMAGE:3936655), partial cds.

DEFINITION BC009760

ACCESSION BC009760

VERSION BC009760.2 GI:33874763

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1815)

AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Ustin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalls, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.

TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

JOURNAL MEDLINE Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

PUBMED 22388257

REFERENCE 2 (bases 1 to 1815)

AUTHORS Strausberg, R.

TITLE Direct Submission

JOURNAL Submitted (02-JUL-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK COMMENT NIH-MGC Project URL: <http://mgc.nci.nih.gov>

On Aug 19, 2003 this sequence version replaced gi:14602504.

Contact: MGC help desk

Email: cgapbs@mail.nih.gov

Tissue Procurement: DCTD/DTP

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Institute for Systems Biology

http://www.systemsbio.org

Contact: amadansystemsbio.org

Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha Madan, Stephanie Rodriguez, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAL Plate: 14 Row: j Column: 14.

Location/Qualifiers

FEATURES
source

1..1815
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3936655"
/tissue_type="Lung, small cell carcinoma"
/clone_lib="NIH MGC 7"
/lab_host="DH10B-R-"
/note="vector: pOTB7"
<1..1815
/gene="C4orf9"
/note="synonym: RES4-25"
/db_xref="LocusID:8602"
<1..1409
/gene="C4orf9"
/codon_start=3
/product="C4orf9 protein"
/protein_id="AAH09760.2"
/db_xref="GI:33874764"
/db_xref="LocusID:8602"

gene

CDS

/translation="CAQREDALELTXLDDDKKEIOTLGHKTPKSNRDKKPKPD
AYDWKRLGEMKQSNRMKTEABLAKBQEHRLKLEARLRMLCKEDENVKXP
KMSADLNDGFDKDDRLLSYKDGQNVVEVDEQEKASDPSPNEEGDSSGG
EDTESPDSHSLSENSEENEPKQRTQPKGLISGERAKGATRDLEPYT
FAAPSEYELSLGSRMERQLLVRIQKNHPSLAEGNKALEKLFGLFLEYVD
LATDDPDLTVIDKLVHLVHLCOMFPESADAIKFLVLDAMHEMEMIETKGRALP
GLDVLKLTGLFPPTSDFWHPVTALVCLSQLTKPTLSLQDVVKGFLVCCFL
EYVALSQRPPELINFLLGILYIATPNKASQSLVHPFRALGNKSNLLVVSAREDA
TWOQSSLSLRWSRIPDLR"
misc_feature
1..1379
/gene="C4orf9"
/note="Nop14; Region: Nop14-like family. Emg1 and Nop14
are novel proteins whose interaction is required for the
maturation of the 18S rRNA and for 40S ribosome
production"
/db_xref="CDD:pfam04147"

misc_feature

ORIGIN

Query Match 1.4%; Score 52; DB 9; Length 1815;
Best Local Similarity 100.0%; Pred. No. 1.8e-14;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3711 TATAATGAAA 3762

DB 1762 TATTAATGAA 1813

RESULT 15

HSM807892
LOCUS Homo sapiens mRNA; cDNA DKFZp686M0187 (from clone DKFZp686M0187).
DEFINITION BX647746
ACCESSION BX647746.1 GI:34366903
VERSION
KEYWORDS
SOURCE
Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 570)
Poustka,A., Albert,R., Moosmayer,P., Schupp,I., Wellenreuther,R.,
Nees,H.W., Weill,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
Wiemann,S.

REFERENCE

AUTHORS
The German Human cDNA Consortium

CONSRM

The German Human cDNA Consortium

TITLE

JOURNAL

COMMENT

Direct Submission
Submitted (27-AUG-2003) MIPS, Ingolstaedter Landstr.1, D-85764
Neuherberg, GERMANY
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
This clone (DKFZp686M0187) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at <http://mips.gsf.de/proj/cDNA/>.

FEATURES

source

1..570
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686M0187"
/tissue_type="human retina"
/clone_lib="686 (synonym: hicc3). Vector pSport1_Sfi; host
DH10B; sites SfiIA + SfiIB"
/dev_stage="adult"
500..505
polyA_signal
polyA_site
511

ORIGIN

Query Match 1.3%; Score 50; DB 9; Length 570;
Best Local Similarity 100.0%; Pred. No. 2e-13;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAA 3762

DB 511 TAAATGAA 560

RESULT 16

BC043545
LOCUS Homo sapiens, clone IMAGE:5171167, mRNA.
DEFINITION BC043545
ACCESSION BC043545.1 GI:27694213
VERSION
KEYWORDS
SOURCE
Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1166)
Strausberg,R.
Direct Submission
Submitted (09-JAN-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REFERENCE

AUTHORS

TITLE

JOURNAL

REMARK

CONTACT: MGC help desk
Email: cgapsb@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcsc.bc.ca

Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saedi, Jacqueline
Schein, Duane Smalhus, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAK Plate: 87 Row: h Column: 16.

FEATURES

Location/Qualifiers
 1..1166
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5171167"
 /tissue type="Brain, adult medulla"
 /clone_lib="NIH MGC_119"
 /lab_host="DH10B"
 /note="Vector: PCMV-SPORT6"

ORIGIN

Query Match 1.3%; Score 50; DB 9; Length 1166;
 Best Local Similarity 100.0%; Pred. No. 1.9e-13;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAA 3762
 |||||
 Db 1031 TAAATGAA 1080

RESULT 17

BC034284 1581 bp mRNA linear PRI 31-JAN-2003
 LOCUS Homo sapiens, clone IMAGE:4779154, mRNA, partial cds.
 DEFINITION BC034284
 ACCESSION BC034284.1 GI:21706848
 KEYWORDS
 SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 1581)
 Strausberg, R.
 Direct Submission
 Submitted (02-JUL-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

REMARK

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT

Contact: MGC help desk
 Email: cgapsb@mail.nih.gov
 Tissue Procurement: James Cleaver, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome
 Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: amg@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,
 Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
 A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAK Plate: 52 Row: d Column: 8.

FEATURES

Location/Qualifiers
 1..1581
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4779154"
 /tissue type="Skin, normal"
 /clone_lib="NCI CGAP_Skn3"
 /lab_host="DH10B"
 /note="Vector: PCMV-SPORT6.ccdB"

CDS

<1..546
 /codon_start=1
 /product="Unknown (protein for IMAGE:4779154)"
 /protein_id="AAH34284.2"

/db_xref="GI:28174969"
 /translation="CARLESPETLKEHVQSDYSLIMCTLOBERVRCRLERQLNDLT
 ELHNEILNLKQELASMEKAYQSYERARIQEALQACQTRISKELQOOQVOVL
 EGLENATNLGKDLINILLAVMVLVSVANCVVPLMKTRRTIFSTLFLVVFIA
 FLWKWDALFSYVERFFSSPR"

ORIGIN

Query Match 1.3%; Score 50; DB 9; Length 1581;
 Best Local Similarity 100.0%; Pred. No. 1.8e-13;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAA 3762
 |||||
 Db 1467 TAAATGAA 1516

RESULT 18

BC008899 1589 bp mRNA linear PRI 03-OCT-2003
 LOCUS Homo sapiens hypothetical protein FLJ23306, mRNA (cdna clone
 MGC:14875 IMAGE:3836842), complete cds.
 DEFINITION BC008899
 ACCESSION BC008899.2 GI:33872024
 KEYWORDS
 SOURCE MGC.

ORGANISM

Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 1589)
 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
 Altshul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, P.,
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
 Stapleton, D., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
 Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
 Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
 Abramson, R.D., Mullaby, S.J., Bosak, S.A., McEwan, P.J.,
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richardson, S.,
 Worley, K.C., Hale, S.S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
 Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Fahey, J., Helton, E., Kettman, M., Madan, A., Rodrigues, S.,
 Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
 Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butterfield, Y.S., Krawinski, M.I., Skalska, U., Smallus, D.E.,
 Schnurch, A., Schein, J.E., Jones, S.J. and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932
 2 (bases 1 to 1589)
 Strausberg, R.

Direct Submission
 Submitted (29-MAY-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 On Aug 19, 2003 this sequence version replaced gi:14286207.

Contact: MGC help desk
 Email: cgapsb@mail.nih.gov
 Tissue Procurement: DCTD/DTP
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
 contact: amadan@systemsbiology.org
 Anup Madan, Jessica Fahey, Erin Helton, Mark Kerteman, Anuradha
 Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAK Plate: 52 Row: d Column: 8.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAK Plate: 52 Row: d Column: 8.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAK Plate: 52 Row: d Column: 8.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAK Plate: 52 Row: d Column: 8.

RESULT 20	BC056324	1639 bp	nrna	linear	VRT 08-OCT-2003
LOCUS	Danio rerio cDNA clone MGC:65884 IMAGE:6800994, complete cds.				
DEFINITION	Danio rerio cDNA clone MGC:65884 IMAGE:6800994, complete cds.				
ACCESSION	BC056324				
VERSION	BC056324.1	GI:33604129			
KEYWORDS	MGC.				
SOURCE	Danio rerio (zebrafish)				
ORGANISM	Danio rerio				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.				
REFERENCE	1 (bases 1 to 1639)				
AUTHORS	Strausberg,R.L., Feingold,B.A., Grouse,L.H., Derge,J.G., Klauener,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altshuler,S.P., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Wax,S.I., Wang,J., Haieh,F., Diatchenko,L., Marusina,K., Farmer,A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldi,M.F., Cabavant,T.I.L., Schetz,T.E., Brownstein,M.J., Usdin,T.B., Tothiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullaly,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettunen,M., Madan,A., Rodriguez,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butlerfield,Y.S., Krzyzanski,M.I., Skalski,U., Smalish,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.				
	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences				
TITLE	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)				
JOURNAL	22388257				
MEDLINE	12477932				
PUBMED	2 (bases 1 to 1639)				
REFERENCE	Strausberg,R.				
AUTHORS	Direct Submission				
TITLE	Submitted (11-AUG-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA				
JOURNAL	NIH-MGC Project URL: http://mgc.nci.nih.gov				
REMARK	Contact: MGC help desk				
COMMENT	Email: cgabs@mail.nih.gov				
	Tissue Procurement: Dr. Chi-Bin Chien				
	cDNA Library Preparation: Invitrogen Corp				
	DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland;				
	Web site: http://www.nisc.nih.gov/				
	Contact: nisc_mgc@hghri.nih.gov				
	Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P., Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legasqui,R., Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C., McDowell,J.J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W., Turgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.				
	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov				
	Series: IRAK Plate: 121 Row: 0 Column: 15				
	This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Similarity but not identity to protein.				
FEATURES	Location/Qualifiers				
source	1. .1639				
	/organism="Danio rerio"				
	/mol_type="mRNA"				


```

/db_host="DH10B"
/note="Vector: pCMV-SPORT6"

Query Match      1.3%; Score 50; DB 9; Length 1654;
Best Local Similarity 100.0%; Pred. No. 1.8e-13; Indels 0; Gaps 0;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY    3713 TAAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
       |||
Db     1578 TAAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1627

RESULT 22
AY069228 Drosophila melanogaster GH27420 full length cDNA. linear INV 17-DEC-2001
LOCUS     AY069228
DEFINITION Drosophila melanogaster (fruit fly)
ACCESSION AY069228.1 GI:17861791
VERSION   FLI_CDNA.
KEYWORDS  Drosophila melanogaster
SOURCE    Drosophila melanogaster
ORGANISM  Drosophila melanogaster [Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.]
REFERENCE 1 (bases 1 to 1737)
AUTHORS   Stapleton,M., Brokstein,P., Hong,L., Agbayani,A., Carlson,J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George,R., Gonzalez,M., Guarin,H., Li.P., Liao,G., Miranda,A., Mungall,C.J., Nunoo,J., Paclet,J., Paragas,V., Park,S., Phouanavong,S., Wan,K., Yu,C., Lewis,S.E., Rubin,G.M. and Celniker,S.
TITLE     Direct Submission
JOURNAL   Submitted (10-DEC-2001) Berkeley Drosophila Genome Project, Lawrence Berkeley National Laboratory, One Cyclotron Road, Berkeley, CA 94720, USA
COMMENT   Sequence submitted by:
          Lawrence Berkeley National Laboratory
          Berkeley, CA 94720
This clone was sequenced as part of a high-throughput process to sequence clones from Drosophila Gene Collection 1 (Rubin et al., Science 2000). The sequence has been subjected to integrity checks for sequence accuracy, presence of a polyA tail and contiguity within 100 kb in the genome. Thus we believe the sequence to reflect accurately this particular cDNA clone. However, there are artifacts associated with the generation of cDNA clones that may have not been detected in our initial analyses such as internal priming, priming from contaminating genomic DNA, retained introns due to reverse transcription of unspliced precursor RNAs, and reverse transcriptase errors that result in single base changes. For further information about this sequence, including its location and relationship to other sequences, please visit our Web site (http://fruitfly.berkeley.edu) or send email to cda@fruitly.berkeley.edu.
FEATURES             source
     Location/Qualifiers
         1..1737
            /organism="Drosophila melanogaster"
            /mol_type="mRNA"
            /strain="y; cn bw sp"
            /db_xref="taxon:7227"
            /map="34A10-34A10"
         1..1737
            /genes="CG9431"
            /note="alignment with genomic scaffold AE003639. gene does not completely overlap longest ORF"
            /db_xref="FLYBASE:FBgn0032484"
            107..1345
               /gene="CG9431"
               /note="Longest ORF"
            /codon_start=1
            /product="GH27420p"
            /protein_id="FAAL39373.1"
            /db_xref="GI:17861792"

ORIGIN

/db_xref="FLYBASE:FBgn0032484"
/translaton="MFRDFVIGNLYTPPTSCCHYPQLGRRLWIEDPPEAPACKPKIL YPTLSTISNTSKENVTLCRVHGSPTNVIAWDYTNYVESRKFSLQRKYIELI REDESKIRKFGHDVFVSRLTVIVNRKSDEGVITCLAEHQPGDSHVHSVVVQDMERI SLIDSNFFAVICLIAMGFMSILFSVLTLIFRKQKPQGHOHTYLQPTSLPVQSPG SEERATAISLLSSGWIRESKIPLDLPSAINEPSKNVTLFKTNSNGSEVMHTFNYSKV RLNSNTYDELNDQAESISRNLKYNLTGNAGDKREKKLKQBELDXSRQSLSQSGNC SRKGQIDEIQDPDLLPSTQPTALKNINETGPSAKAEVNPRNKINGQTLYKEKX SVRIKNIISTKEPTGTVDISI"

Query Match      1.3%; Score 50; DB 3; Length 1737;
Best Local Similarity 100.0%; Pred. No. 1.8e-13; Indels 0; Gaps 0;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY    3713 TAAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
       |||
Db     1662 TAAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1711

RESULT 23
BC053205 Dario rerio cDNA clone MGC:64021 IMAGE:6793127, complete cds. linear VRT 07-OCT-2003
LOCUS     BC053205
DEFINITION Dario rerio cDNA clone MGC:64021 IMAGE:6793127, complete cds.
ACCESSION BC053205
VERSION   BC053205.1 GI:31418862
KEYWORDS  MGC.
SOURCE    Dario rerio (zebrafish)
ORGANISM  Dario rerio
REFERENCE 1 (bases 1 to 1786)
AUTHORS   Strausberg R.L., Feingold, B.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeng, K.H., White, O., Brown, A.P., Holt, I.A., Kane, J.S., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Weihe, F., Dietzschko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stabileton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshyuki, S., Carninci, P., Prange, C., Raja, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Wille, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickinson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalios, D.E., Schermer, A., Schein, J.E., Jones, S.J. and Warra, M.A.
TITLE     Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
MEDLINE   22398257
PubMed    12477932
2 (bases 1 to 1786)
Strausberg, R.
Direct Submission
TITLE     Submitted (02-JUN-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT   Contact: MGC help desk
           Email: gcapbs@mail.nih.gov
           Tissue Procurement: Leonard I. Zon, M.D.
           CDNA Library Preparation: Invitrogen Corp
           CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLNL)
           DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
           Web site: http://www.shgc.stanford.edu
           Contact: (Dickson, Mark) mcdgpaxil.stanford.edu
```


Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, J., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAK Plate: 117 Row: g Column: 24
 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, similarity but not identity to protein.

FEATURES

source

1. 1786
 /organism="Danio rerio"
 /mol_type="mRNA"
 /db_xref="taxon:7955"
 /clone="MGC:64021 IMAGE:6793127"
 /tissue_type="kidney, zebrafish"
 /clone_lib="NCI CGAP_ZKId1"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6.1"

CDS

313..1545
 /codon_start=1
 /product="Unknown (protein for MGC:64021)"
 /protein_id="AAH53205.1"
 /db_xref="GI:31418863"
 /translation="MVLISQRDELNRAIADYLRNGYEAYVFKKEALDMNEELD
 KYAGLLEKWTIVIRLQKKMELESINIGGIGQKRPKEWIPRPPEKY
 ALSGRSPVTRVIFHPVSVTVSASEDATIKWDHETGDFERTLKGTDSVODISFDH
 TGKLLASCSADMTIKLWDROGFCIRTWGHGHNVSVAIMPNGDHTVSAASRDKTIKM
 HEVATGCVKFTGHRNVMVRNQGTLIASSNDQTVRVVVAATKECAELREHE
 HVVESCISWAPESAPHTLEATGSETKSGKGFPELLSGDKTIKMDVSGWGLMTIL
 VGHDNVGVGLVHPGGKIVTSACDDKTLIRINDYKNKRCITKLSAHEHFVTSLDFPKHTA
 PYYVTGSDVDQTVKWECR"

misc_feature

331..426
 /note="L18; Region: Lissencephaly type-1-like homology motif"

misc_feature

601..1542
 /note="COG2319; Region: FOG: WD40 repeat [General function prediction only]"
 /db_xref="CDD:COG2319"

ORIGIN

Query Match 1.3%; Score 50; DB 5; Length 1786;

Best Local Similarity 100.0%; Pred. No. 1.8e-13;

Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAAAA 3762

Db 1731 TAAATGAAAAA 1780

RESULT 24

AK026959

LOCUS Homo sapiens cDNA: FLJ23306 fis, clone HEP11541. linear PRI 12-SEP-2003

DEFINITION AK026959

ACCESSION AK026959

VERSION AK026959.1 GI:10439943

KEYWORDS oligo capping; fis (full insert sequence).

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 Kawabata, A., Hikiji, T., Kobatake, N., Inagaki, H., Ikema, Y.,

Okamoto, S., Okitani, R., Ota, T., Suzuki, Y., Obayashi, M., Nishi, T.,

Shibahara, T., Tanaka, T., Nakamura, Y., Isogai, T. and Sugano, S.

NEDO human cDNA sequencing project

Unpublished

2 (bases 1 to 1820)

Sugano, S., Suzuki, Y., Ota, T., Obayashi, M., Nishi, T., Isogai, T.,

Shibahara, T., Tanaka, T. and Nakamura, Y.

Direct Submission

Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,

University of Tokyo, Laboratory of Genome Structure Analysis, Human Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail: fcdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286, Fax:81-3-5449-5416)

COMMENT NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing; Department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).

FEATURES

source

1. 1820
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="HEP11541"
 /cell_line="HepG2"
 /cell_type="hepatoma"
 /clone_lib="HSP"
 /note="cloning vector pME18SFL3"

CDS

563..931
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="BAB15603.1"
 /db_xref="GI:10439944"
 /translation="MSVGLDLPQLPGSPSPSLKEAFADGEAGEGGGRPSLEWRLOQ
 LLPOPLSLSLTSPQGTGPFSLPWCILDQVLSLLHSPVTPVHSSPGSKQAV
 NSWPELSLWLVAHAPFLVVC"

ORIGIN

Query Match 1.3%; Score 50; DB 9; Length 1820;

Best Local Similarity 100.0%; Pred. No. 1.8e-13;

Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAAAA 3762

Db 1719 TAAATGAAAAA 1768

RESULT 25

BC052972

LOCUS Danio rerio cDNA clone MGC:56344 IMAGE:5604031, complete cds. linear VRT 07-OCT-2003

DEFINITION BC052972

ACCESSION BC052972.1 GI:31418214

KEYWORDS MGC.

SOURCE Danio rerio (zebrafish)

ORGANISM Danio rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Ostariophysi;

Cypriniformes; Cyprinidae; Danio.

1 (bases 1 to 1880)

Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,

Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,

Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,

Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,

Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,

Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,

Schetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,

Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,

Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,

McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,

Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,

Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,

Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y.,

Bouffard, G.C., Blakesley, R.W., Touchman, O.W., Green, E.D.,

Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,

Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalls, D.E.,

Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length

human and mouse cDNA sequences

TITLE

JOURNAL

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 22388257
 12477932
 2 (bases 1 to 1880)
 Strausberg,R.
 Direct Submission
 Submitted (02-JUN-2003) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Dr. Sumio Sugano
 cDNA Library Preparation: Dr. Sumio Sugano
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www-shgc.stanford.edu>
 Contact: (Dickson, Mark) mdc@paxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAK Plate: 103 Row: P Column: 5
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis, Similarity but not identity to protein.

FEATURES
 source
 1..1880
 /organism="Danio rerio"
 /mol_type="mRNA"
 /db_xref="taxon:7955"
 /clone="MGC:56344 IMAGE:5604031"
 /tissue_type="Whole body, adult male"
 /clone_lib="Sugano SJF adult male"
 /lab_host="DH10B"
 /note="vector: pME18S-FL3"
 227..703
 /codon_start=1
 /product="Unknown (protein for MGC:56344)"
 /protein_id="AAH52972.1"
 /db_xref="GI:131418215"
 /translation="MPGKGKKGKGRGKQFSPNPEIDRQWRAQKELENGGAERGSS
 ESEETSSDDEQPKKVEGLIETENPNVSKVKEIDVNAPELRREREIEK
 QKAKERYMKLHLEKTEQARADLARLAIKKQEDRAKKKEELRKEKEAEKSKR"

CDS
 227..703
 /codon_start=1
 /product="Unknown (protein for MGC:56344)"
 /protein_id="AAH52972.1"
 /db_xref="GI:131418215"

ORIGIN
 Query Match 1.3%; Score 50; DB 5; Length 1880;
 Best Local Similarity 100.0%; Pred. No. 1.8e-13;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 3713 TAAATGAAA 3762
 Db 1817 TAAATGAAA 1866

RESULT 26
 BC036869
 LOCUS 2275 bp mRNA linear PRI 07-OCT-2003
 DEFINITION Homo sapiens hypothetical protein LOC283130, mRNA (cDNA clone
 MGC:46503 IMAGE:5228549), complete cds.
 ACCESSION BC036869
 VERSION BC036869.1 GI:22477390
 KEYWORDS MGC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 2275)
 Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
 Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,

Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
 Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Heien,F.,
 Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
 Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
 Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
 Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
 Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
 McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
 Worley,K.C., Hale,S., Garcia,A.M., Gay,D.J., Hulyk,S.W.,
 Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
 Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,A.,
 Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
 Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
 Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
 Schnerch,A., Schein,J.E., Jones,S.J., Skalske,U., Smailus,D.E.,
 human and mouse cDNA sequences
 Generation and initial analysis of more than 15,000 full-length
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 22388257
 12477932
 2 (bases 1 to 2275)
 Strausberg,R.
 Direct Submission
 Submitted (23-AUG-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC),
 Gaithersburg, Maryland;
 Web site: <http://www.nisc.nih.gov/>
 Contact: nisc_mgc@nigr.nih.gov
 Akhter,N., Ayale,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
 Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
 Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
 Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
 Maduro,Q.L., Mastello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C.,
 McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,
 Tsurgouon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
 Young,A., Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAK Plate: 78 Row: H Column: 3.
 Location/Qualifiers
 1..2275
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:46503 IMAGE:5228549"
 /tissue_type="Pancreas, Spleen, adult pooled"
 /clone_lib="NIH MGC_120"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 1..2275
 /gene="LOC283130"
 /db_xref="LocusID:283130"
 242..982
 /codon_start=1
 /product="LOC283130 protein"
 /protein_id="AAH36869.1"
 /db_xref="GI:22477391"
 /db_xref="LocusID:283130"
 /translation="WVKYRHSLLGFFKMGSPFIASVNVSLFVGVSTLLVLVLA
 TSHQRRRAQPPSYMHIPLAGCTGGGLQALCPDLIKVRLQNTQEPRAQSPPPRY
 QQPVHCAISIFREGPRGLFRGAWALTLDITFTVIGIYTYEGLCRQITPEQGNFSSA

FEATURES
 source
 1..2275
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:46503 IMAGE:5228549"
 /tissue_type="Pancreas, Spleen, adult pooled"
 /clone_lib="NIH MGC_120"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 1..2275
 /gene="LOC283130"
 /db_xref="LocusID:283130"
 242..982
 /codon_start=1
 /product="LOC283130 protein"
 /protein_id="AAH36869.1"
 /db_xref="GI:22477391"
 /db_xref="LocusID:283130"
 /translation="WVKYRHSLLGFFKMGSPFIASVNVSLFVGVSTLLVLVLA
 TSHQRRRAQPPSYMHIPLAGCTGGGLQALCPDLIKVRLQNTQEPRAQSPPPRY
 QQPVHCAISIFREGPRGLFRGAWALTLDITFTVIGIYTYEGLCRQITPEQGNFSSA

Worley, K.C. Direct Submission Submitted (27-AUG-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA 4 (bases 1 to 96967) Worley, K.C. Direct Submission Submitted (25-SEP-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA 5 (bases 1 to 96967) Worley, K.C. Direct Submission Submitted (28-FEB-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA On Sep 25, 2002 this sequence version replaced gi:22380679. INFORMATION: http://www.hgsc.bcm.tmc.edu/ or email gc-help@bcm.tmc.edu	Worley, K.C. Direct Submission Submitted (27-AUG-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA 4 (bases 1 to 96967) Worley, K.C. Direct Submission Submitted (25-SEP-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA 5 (bases 1 to 96967) Worley, K.C. Direct Submission Submitted (28-FEB-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA On Sep 25, 2002 this sequence version replaced gi:22380679. INFORMATION: http://www.hgsc.bcm.tmc.edu/ or email gc-help@bcm.tmc.edu
CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.	CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.
ANNOTATION OF FEATURES: STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts. Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences. Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.	ANNOTATION OF FEATURES: STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts. Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences. Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.
SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.	SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.
QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: http://www.hgsc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html .	QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: http://www.hgsc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html .
FEATURES source 1..96967 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /chromosome="12" /clone="RP11-324P9" complement(1..2004) /notes="overlaps bases 1..2004 of clone AC139931" /function="clone overlap" 55964..56062 /standard_name="D11S3430" 58024..58294 /standard_name="SHGC-111995" 58082..58416 /standard_name="SHGC-145133" 50409..50681 /standard_name="SHGC-144174"	FEATURES source 1..96967 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /chromosome="12" /clone="RP11-324P9" complement(1..2004) /notes="overlaps bases 1..2004 of clone AC139931" /function="clone overlap" 55964..56062 /standard_name="D11S3430" 58024..58294 /standard_name="SHGC-111995" 58082..58416 /standard_name="SHGC-145133" 50409..50681 /standard_name="SHGC-144174"

Query Match	1.3%;	Score 50;	DB 9;	Length 3606;
Best Local Similarity	100.0%;	Pred. No. 1.7e-13;		
Matches 50;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	3713	TAAATGAAA 3762		
Db	3531	TAAATGAAA 3580		
RESULT 29				
AC127894				
LOCUS	Homo sapiens 12 BAC RP11-324P9 (Roswell Park Cancer Institute Human			
DEFINITION	BAC library) complete sequence.			
ACCESSION	AC127894			
VERSION	AC127894.6			
KEYWORDS	HTG.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	AC127894 96967 bp DNA linear PRI 28-FEB-2003			
AUTHORS	Worley, K.C.			
TITLE	Submitted (28-FEB-2003) Human Genome Sequencing Center, Department			
JOURNAL	of Molecular and Human Genetics, Baylor College of Medicine, One			
REFERENCE	Baylor Plaza, Houston, TX 77030, USA			
Unpublished	2 (bases 1 to 96967)			
Worley, K.C.	Direct Submission			
Submitted (19-JUL-2001)	Human Genome Sequencing Center, Department			
of Molecular and Human Genetics, Baylor College of Medicine, One	Baylor Plaza, Houston, TX 77030, USA			
3 (bases 1 to 96967)				

```

STS      87156. .87296
         /standard_name="RH104171"
misc_feature 91175. .92742
         /note="imperfect tandem repeat, confirmed by restriction
         map and PCR"
misc_feature 94964. .96967
         /function="unresolved tandem repeat"
         /note="overlaps bases 1. .2004 of clone AC020656"
         /function="clone overlap"

ORIGIN
Query Match      1.3%; Score 50; DB 9; Length 96967;
Best Local Similarity 100.0%; Pred. No. 1.2e-13;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2559 CATTCCAGCTGGGCAACAGAGTGAGACCCCTGCTCAAAAAAAAAAAAAA 2608
      |||
Db 38683 CATTCCAGCTGGGCAACAGAGTGAGACCCCTGCTCAAAAAAAAAAAAAA 38732

RESULT 30
AR425685
LOCUS      AR425685      159 bp      DNA      linear      PAT 18-DEC-2003
DEFINITION Sequence 17182 from patent US 6639063.
ACCESSION AR425685
VERSION   AR425685.1 GI:40180795
KEYWORDS  Unknown.
SOURCE    Unknown.
ORGANISM  Unclassified.
REFERENCE 1 (bases 1 to 159)
AUTHORS  Edwards,J.-B.D.M., Jobert,S. and Giordano,J.-Y.
TITLE    EST's and encoded human proteins
JOURNAL  Patent: US 6639063-A 17182 28-OCT-2003;
FEATURES  Location/Qualifiers
          source
            1. .159
            /organism="unknown"
            /mol_type="genomic DNA"

ORIGIN
Query Match      1.3%; Score 49; DB 6; Length 159;
Best Local Similarity 100.0%; Pred. No. 7.2e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3714 AAATGAAAAA 3762
      |||
Db 100 AAATGAAAAA 148

RESULT 31
BD121238
LOCUS      BD121238      159 bp      DNA      linear      PAT 18-SEP-2002
DEFINITION EST and encoded human protein.
ACCESSION BD121238
VERSION   BD121238.1 GI:23216148
KEYWORDS  JP 2002010789-A/13315.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
          Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 159)
AUTHORS  Edwards,J.B.D.M., Jobert,S. and Giordano,J.E.
TITLE    EST and encoded human protein
JOURNAL  Patent: JP 2002010789-A 13315 15-JAN-2002;
          GENSET CORP
COMMENT   OS Homo sapiens (human)
          PN JP 2002010789-A/13315
          PD 15-JAN-2002
          PF 07-AUG-2000 JP 2002080989
          PR 05-AUG-1999 US 60/147499
          PI JEAN BAPTISTE DOMAS MILNE EDWARDS,SEVELIN JOBERT,JEAN EVE PI
          GIORDANO
          PC C12N15/09,C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19, PC

Qy 3714 AAATGAAAAA 3762
      |||
Db 100 AAATGAAAAA 148

RESULT 32
BC038545
LOCUS      BC038545      444 bp      mRNA      linear      PRI 15-OCT-2002
DEFINITION Homo sapiens, clone IMAGE:5224224, mRNA.
ACCESSION BC038545
VERSION   BC038545.1 GI:23959014
KEYWORDS
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 444)
AUTHORS  Strausberg,R.
TITLE    Direct Submission
JOURNAL  Submitted (09-OCT-2002) National Institutes of Health, Mammalian
          Gene Collection (MGC), Cancer Genomics Office, National Cancer
          Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
          USA
          NIH-MGC Project URL: http://mgc.nci.nih.gov
          Contact: MGC help desk
          Email: cgapbs-r@mail.nih.gov
          Tissue Procurement: Life Technologies, Inc.
          cDNA Library Preparation: Life Technologies, Inc.
          cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNA Sequencing by: Genome Sequence Centre,
          EC Cancer Agency, Vancouver, BC, Canada
          info@cgsc.bc.ca
          Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
          Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
          Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
          Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
          Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
          Schein, Duane Smalus, Michael Smith, Lorraine Spence, Jeff Stott,
          Michael Thorne, Miranada Teai, Natasja van den Bosch, Jill Vardy,
          George Yang, Scott Zuyderduyn, Marco Marra.

REMARK
COMMENT   Clone distribution: MGC clone distribution information can be found
          through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
          Series: IRAC Plate: 87 Row: m Column: 12
          This clone was selected for full length sequencing because it
          passed the following selection criteria: UCSC EST-genomic
          comparison.

FEATURES
          Location/Qualifiers
            1. .444
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:5224224"
            /tissue_type="Pancreas, Spleen, adult pooled"
            /clone_lib="NIH_MGC_120"

```

```

C12N1/21,
PC C12N5/10,C12P21/02,C12P21/08,C12Q1/68,C12N15/00,C12N5/00, PC
C12N15/00
CC EST and encoded human protein
FH Key Location/Qualifiers
FT source 1. .159
FT /organism="Homo sapiens (human)"
FEATURES
          Location/Qualifiers
            1. .159
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"

ORIGIN
Query Match      1.3%; Score 49; DB 6; Length 159;
Best Local Similarity 100.0%; Pred. No. 7.2e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3714 AAATGAAAAA 3762
      |||
Db 100 AAATGAAAAA 148

RESULT 32
BC038545
LOCUS      BC038545      444 bp      mRNA      linear      PRI 15-OCT-2002
DEFINITION Homo sapiens, clone IMAGE:5224224, mRNA.
ACCESSION BC038545
VERSION   BC038545.1 GI:23959014
KEYWORDS
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 444)
AUTHORS  Strausberg,R.
TITLE    Direct Submission
JOURNAL  Submitted (09-OCT-2002) National Institutes of Health, Mammalian
          Gene Collection (MGC), Cancer Genomics Office, National Cancer
          Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
          USA
          NIH-MGC Project URL: http://mgc.nci.nih.gov
          Contact: MGC help desk
          Email: cgapbs-r@mail.nih.gov
          Tissue Procurement: Life Technologies, Inc.
          cDNA Library Preparation: Life Technologies, Inc.
          cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNA Sequencing by: Genome Sequence Centre,
          EC Cancer Agency, Vancouver, BC, Canada
          info@cgsc.bc.ca
          Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
          Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
          Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
          Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
          Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
          Schein, Duane Smalus, Michael Smith, Lorraine Spence, Jeff Stott,
          Michael Thorne, Miranada Teai, Natasja van den Bosch, Jill Vardy,
          George Yang, Scott Zuyderduyn, Marco Marra.

REMARK
COMMENT   Clone distribution: MGC clone distribution information can be found
          through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
          Series: IRAC Plate: 87 Row: m Column: 12
          This clone was selected for full length sequencing because it
          passed the following selection criteria: UCSC EST-genomic
          comparison.

FEATURES
          Location/Qualifiers
            1. .444
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:5224224"
            /tissue_type="Pancreas, Spleen, adult pooled"
            /clone_lib="NIH_MGC_120"

```



```

3714 AAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
|||||
650 AAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 698

RESULT 35
BC060283 752 bp mRNA linear ROD 21-OCT-2003
DEFINITION Mus musculus cDNA clone IMAGE:30291051, partial cds.
ACCESSION BC060283
VERSION BC060283.1 GI:37805219
KEYWORDS Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 752)
Straussberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, J., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B.S., Buetow, K.H., Schaefer, C.F., Shat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.P., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Udwin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullen, J.S., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, D.J., Hulyk, S.W.,
Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahney, J., Helton, E., Kettner, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzyzanski, M.I., Skalska, U., Smalios, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
12477932
2 (bases 1 to 752)
Straussberg, R.
Direct Submission
Submitted (20-OCT-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabbs@mail.nih.gov
Tissue Procurement: Dr. David Rowe
cDNA Library Preparation: Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzyzanski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pardoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
Schein, Duane Smalios, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC plate: 131 Row: e Column: 17
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 13385879.
Location/Qualifiers
1. .752

/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:30291051"
/tissue_type="Jaw and Limb, mouse, day 18.5 to new born,
pool of mature fore and hind limb, maxilla and mandible
containing endochondral and membranous bone, formed
joints, tendon, ligaments, dermis, epidermis, muscle and
teeth with newly forming dentin and enamel"
/clone_lib="NIH MGC_136"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6.1"

ORIGIN
Query Match 1.3%; Score 49; DB 10; Length 752;
Best Local Similarity 100.0%; Pred. No. 6.2e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
|||||
DB 697 AAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 745
|||||

RESULT 36
S71381 959 bp mRNA linear PRI 01-OCT-1994
LOCUS prosome beta-subunit-multicatalytic proteinase complex [human,
lymphoma, mRNA, 959 nt].
DEFINITION
ACCESSION S71381
VERSION S71381.1 GI:551546
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 959)
Gerards, W.L., Hop, F.W., Hendriks, I.L. and Bloemendal, H.
Cloning and expression of a human pro(tea)some beta-subunit cDNA: a
homologue of the yeast PRE4-subunit essential for
peptidylglutamyl-peptide hydrolase activity
FEBS Lett. 346 (2-3), 151-155 (1994)
94283586
8013624
JOURNAL
MEDLINE
PUBMED
REMARK GenBank staff at the National Library of Medicine created this
entry [NCBI gi354150354] from the original journal article.
This sequence comes from Fig. 1.
FEATURES
Location/Qualifiers
1. .959
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1. .959
/genes="prosome beta-subunit, HSBpros26"
100. .801
/genes="prosome beta-subunit, HSBpros26"
/notes="multicatalytic proteinase complex; Yeast
PRE4-subunit homolog; Method: conceptual translation with
partial peptide sequencing; This sequence comes from Fig.
1; HSBpros26"
/codon_start=1
/product="prosome beta-subunit"
/protein_id="AAB31085.1"
/db_xref="GI:551547"
/tranlation="MDPASALYRGPTTRQNTQNTGTSVLGVKFGGVVIAADMLGSY
GSLARFNISIRNVRNNSYMLGASGDYADFOYLKQVLQGVVIDEILGDSHSYSPRAI
HSWLTIRYRSRKSQNPNTWNTVIGYADGESFLGVDMVLGVAYEARPSLATGYGALA
QPLLRLVLEKQPVLSQTEARDLVERCMVLIYRDARSYNRFTATVTEKGEVEIGPLS
TETNWDIAHMSIGFE"

ORIGIN
Query Match 1.3%; Score 49; DB 9; Length 959;
Best Local Similarity 100.0%; Pred. No. 6e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


```


Qy 3714 AAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Dd 904 AAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 952

RESULT 37					
BC032451					
LOCUS	BC032451	997 bp	mRNA	linear	PRI 19-NOV-2003
DEFINITION	Homo sapiens cDNA clone MGC:40426 IMAGE:5178085, complete cds.				

ACCESSION	BC032451
VERSION	BC032451.1
	GI:21619605

KEYWORDS	MGC.	Homosapiens (human)
SOURCE		

ORGANISM Homo sapiens

Mammalia; Eutheria;

AUTHORS
Strausberg, R.L., Feid

Klausner, R.D., Collins
Altschul, S.F., Zeebe

Hopkins, R. F., Jordan
Diatchenko I. Maris

Stapleton, M., Soares

Carninci, P., Prange,

McKernan, K. J., Malek

Worley, K.C., Hale, S.
Villalon, D.K., Muzny

Fahey, J., Helton, E.,
Sanchez A. Whiting

Bouffard, G.G., Blake

Butterfield, Y.S., Kriz

TITLE	SCHEMCH, A., SCHEMCH, A.
Generation and initiation	

JOURNAL OF THE
HUMAN AND MOUSE CDNA
PROC. NATL. ACAD. SCI.

MEDLINE 22388257
 PUBMED 12477932

REFERENCE 2 (bases 1 to 997)

TITLE	Direct Submission
1. <u>THE</u>	
2. <u>THE</u>	
3. <u>THE</u>	
4. <u>THE</u>	
5. <u>THE</u>	
6. <u>THE</u>	
7. <u>THE</u>	
8. <u>THE</u>	
9. <u>THE</u>	
10. <u>THE</u>	
11. <u>THE</u>	
12. <u>THE</u>	
13. <u>THE</u>	
14. <u>THE</u>	
15. <u>THE</u>	
16. <u>THE</u>	
17. <u>THE</u>	
18. <u>THE</u>	
19. <u>THE</u>	
20. <u>THE</u>	
21. <u>THE</u>	
22. <u>THE</u>	
23. <u>THE</u>	
24. <u>THE</u>	
25. <u>THE</u>	
26. <u>THE</u>	
27. <u>THE</u>	
28. <u>THE</u>	
29. <u>THE</u>	
30. <u>THE</u>	
31. <u>THE</u>	
32. <u>THE</u>	
33. <u>THE</u>	
34. <u>THE</u>	
35. <u>THE</u>	
36. <u>THE</u>	
37. <u>THE</u>	
38. <u>THE</u>	
39. <u>THE</u>	
40. <u>THE</u>	
41. <u>THE</u>	
42. <u>THE</u>	
43. <u>THE</u>	
44. <u>THE</u>	
45. <u>THE</u>	
46. <u>THE</u>	
47. <u>THE</u>	
48. <u>THE</u>	
49. <u>THE</u>	
50. <u>THE</u>	
51. <u>THE</u>	
52. <u>THE</u>	
53. <u>THE</u>	
54. <u>THE</u>	
55. <u>THE</u>	
56. <u>THE</u>	
57. <u>THE</u>	
58. <u>THE</u>	
59. <u>THE</u>	
60. <u>THE</u>	
61. <u>THE</u>	
62. <u>THE</u>	
63. <u>THE</u>	
64. <u>THE</u>	
65. <u>THE</u>	
66. <u>THE</u>	
67. <u>THE</u>	
68. <u>THE</u>	
69. <u>THE</u>	
70. <u>THE</u>	
71. <u>THE</u>	
72. <u>THE</u>	
73. <u>THE</u>	
74. <u>THE</u>	
75. <u>THE</u>	
76. <u>THE</u>	
77. <u>THE</u>	
78. <u>THE</u>	
79. <u>THE</u>	
80. <u>THE</u>	
81. <u>THE</u>	
82. <u>THE</u>	
83. <u>THE</u>	
84. <u>THE</u>	
85. <u>THE</u>	
86. <u>THE</u>	
87. <u>THE</u>	
88. <u>THE</u>	
89. <u>THE</u>	
90. <u>THE</u>	
91. <u>THE</u>	
92. <u>THE</u>	
93. <u>THE</u>	
94. <u>THE</u>	
95. <u>THE</u>	
96. <u>THE</u>	
97. <u>THE</u>	
98. <u>THE</u>	
99. <u>THE</u>	
100. <u>THE</u>	

JOURNAL
SUBMITTED (06-JUN-2006)
Gene Collection (MGCC)

Institute, 31 Center
USA

REMARK
NIH-MGC Project URL: <http://www.genome.gov/27532711>

Email: cgapbs-r@mail.ru

CDNA Library Preparation

CDNA Library Arrayed
DNA Sequencing by: N

Sequencing Center (Marjorie Catthersburg Maryland)

Web site: <http://www.>

Akhter, N., Ayele, K.,

Brakesley, R.W., Boulder
Dietrich, N.L., Grants

Hansen, N., Ho, S.-L.,
Maduro, O.L., Masiel-

McDowell, J., Pearson
Tsurgren C. Vogt J.

Young, A., Zhang, L. - H

Clone distribution:

Series: IRAK Plate:

This clone was selected and passed the following

```

FEATURES             Location/Qualifiers
     source            1..397
                        /organism="Homo sapiens"
                        /mol_type="mRNA"
                        /db_xref="taxon:9606"
                        /clone="MGC:40426 IMAGE:5178085"
                        /issue_type="Brain, Lung, Testis, adult, pooled whole"
                        /clone_lib="NIH MGC_115"
                        /lab_host="DH10B"
                        /note="Vector: pCMV-SPORT6"
     CDS               14..724
                        /codon_start=1
                        /product="Unknown (protein for MGC:40426)"
                        /protein_id="AAH32451.1"
                        /db_xref="GI:21619606"
                        /translation="MTEPQAQLFLLILLMLPDSNGENVLTQSPCTLSLSPGERATLSCR
                        ACSUSSSLXAWYQKPGQAPRLIIYGVSSRATGIFDRSGSGSGTDFLTITSLRPER
                        DFPAVYCYQGYQSRPTFGQGRKLDIKRTVAAPSVFIFPPSDEQLSKASVAVCLLNN
                        FYPRAKQVMQVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEV
                        THQGLSSPVTKSFNRGEC"
     misc_feature      104..400
                        /note="IGV; Region: Immunoglobulin domain variable region
                        (v) subfamily"
                        /db_xref="CDD:cd00099"
     misc_feature      407..709
                        /note="IGC; Region: Immunoglobulin domain constant region
                        subfamily"
                        /db_xref="CDD:cd00098"

ORIGIN
Query Match          1.3%; Score 49; DB 9; Length 997;
Best Local Similarity 100.0%; Pred. No. 6e-13; 0; Indels 0; Gaps 0;
Matches 49; Conservative 0; Mismatches 0;

QY 3714 AAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
      |||
Db 920 AAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 968

RESULT 38
BC000761
LOCUS
DEFINITION
Homo sapiens SNARE associated protein snapin, mRNA (cdna clone
MGC:2717 IMAGE:2821705), complete cds.
ACCESSION
BC000761
VERSION
BC000761.2 GI:38014219
KEYWORDS
MGC.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1008)
Srausberg,R.L., Reingold,E.A., Grouse,L.H., Derge,J.G.
Klauser,R.D., Collins,F.S., Wagner,L., Shennan,C.M., Schuler,G.D.,
Altschul,S.F., Zerbe,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farner,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raja,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullany,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.C., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Vallalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettman,M., Madan,A., Rodrigues,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Buffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,F.S., Krzywinski,M.I., Skalska,U., Smalls,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marr,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences

```


/db_xref="LocusID:28981"
 /translation="MKNLEVLQRLATDEMKAIVSSDQQRKKAIREQYTKNTARQEN
 LKGLREKQKVIRESHGPNKQAKWRDLQLECKQCFKQSQTSIQGVQEGGE
 DRLLIL"

ORIGIN
 Query Match 1.3%; Score 49; DB 9; Length 1030;
 Best Local Similarity 100.0%; Pred. No. 6e-13;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3714 AAATCAA 3762

Db 948 AAATGAAA 996

RESULT 40
 BC038410 1051 bp mRNA linear PRI 06-OCT-2003
 LOCUS Homo sapiens LOC92346, mRNA (cDNA clone MGC:35143 IMAGE:5169137),
 complete cds.

ACCESSION BC038410
 VERSION BC038410.1 GI:23468237
 KEYWORDS MGC.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1051)

Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,

Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,

Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,

Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,

Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,

Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,

Schetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,

Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,

Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,

McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,

Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,

Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,

Fahy, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,

Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,

Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,

Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,

Butterfield, Y.S., Krzywinski, M.I., Skalka, U., Smalhus, D.E.,

Schmerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length

human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

22388257

12477932

2 (bases 1 to 1051)

Strausberg, R.

Direct Submission

Submitted (01-OCT-2002) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 1A03, Bethesda, MD 20892-2590,

USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgabs@mail.nih.gov

Tissue procurement: Life Technologies, Inc.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Baylor College of Medicine Human Genome

Sequencing Center

Center code: BCM-HGSC

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>

Contact: ang@bcm.tmc.edu

Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louieged, H.,

Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,

A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAX Plate: 51 Row: b Column: 5
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 21321731.

FEATURES

source

1..1051

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:35143 IMAGE:5169137"

/tissue type="Brain, adult medulla"

/clone_lib="NIH MGC_119"

/lab_host="DH10B"

/notes="Vector: pCMV-SPORT6"

1..1051

/genes="LOC92346"

/db_xref="LocusID:92346"

197..748

/codon_start=1

/product="LOC92346"

/protein_id="AA383410.1"

/db_xref="GI:23468238"

/db_xref="LocusID:92346"

/translation="MEKRELKASVPKFDKIPWLSEASLVNKPVLVSLPRRYPHTSATF
 LTSSKKNMLPILFQVPDVLKARNQCDMLRNQQLCSQCEQKMKVQPRTKIPDD
 PKASFNCMSYRMSLHOPKFPQTPPEFDDIPTENHYRLPLIGPRTAVFHGLITAY
 KTLKERQSRSLPRKEFIGTKTRQ"

ORIGIN

Query Match 1.3%; Score 49; DB 9; Length 1051;

Best Local Similarity 100.0%; Pred. No. 6e-13;

Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3714 AAATGAAA 3762

Db 976 AAATGAAA 1024

RESULT 41

BC004905

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

AUTHORS

TITLE

JOURNAL

REMARK

COMMENT

BC004905 1068 bp mRNA linear PRI 16-SEP-2003
 IMAGE:3509565), complete cds.

BC004905.2 GI:33872888

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1068)

Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,

Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,

Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,

Hopkins, R.E., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,

Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,

Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,

Schetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,

Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,

Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,

McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,

Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,

Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,

Fahy, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,

Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,

Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,

Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,

Butterfield, Y.S., Krzywinski, M.I., Skalka, U., Smalhus, D.E.,

Schmerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length

human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

```

MEDLINE      22388257
PUBMED       12477932
REFERENCE    2 (bases 1 to 1068)
AUTHORS      Strausberg,R.
TITLE        Direct Submission
JOURNAL      Submitted (21-MAR-2001) National Institutes of Health, Mammalian
              Gene Collection (MGC), Cancer Genomics Office, National Cancer
              Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
              USA
REMARK       NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT       On Aug 19, 2003 this sequence version replaced gi:13436187.
              Contact: MGC help desk
              Email: cgapbs-r@mail.nih.gov
              Tissue Procurement: ATCC
              cDNA Library Preparation: Rubin Laboratory
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
              DNA Sequencing by: Institute for Systems Biology
              http://www.systemsbio.org
              Contact: amadan@systemsbiology.org
              Anup Madan, Jessica Fahay, Erin Helton, Mark Kettelman, Anuradha
              Madan, Stephanie Rodriguez, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 11 Row: 6 Column: 7
This clone has the following problem: The cds is short compared to
the longest cds in the locus.

FEATURES             source
1..1068
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3509565"
/tissue_type="Brain, neuroblastoma"
/clone_lib="NIH MGC 19"
/lab_host="DH10B-R"
/notes="Vector: pOTB7"
1..1068
/gene="MRP52"
/notes="synonyms: CGI-91, MRP-S2"
/db_xref="LocusID:51116"
291..452
/codon_start=1
/product="MRP52 protein"
/protein_id="AAH04905.2"
/db_xref="GI:33872889"
/db_xref="LocusID:51116"
/translation="MTTLRLWLCVSTAGSSRRSPGPRRSGSLRLSIACRRRSPGTR
GQPTLLGLT"

ORIGIN
Query Match      1.3%; Score 49; DB 9; Length 1068;
Best Local Similarity 100.0%; Pred. No. 6e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
    |||||
Db 999 AAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1047

RESULT 42
BC063150
LOCUS          Rattus norvegicus cDNA clone MGC:72699 IMAGE:6922267, complete cds.
DEFINITION
ACCESSION      BC063150
VERSION        BC063150.1 GI:38649091
KEYWORDS       MGC.
SOURCE          Rattus norvegicus (Norway rat)
ORGANISM       Rattus norvegicus
                Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
                Rattus.
REFERENCE      1 (bases 1 to 1109)

```

```

AUTHORS      Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
              Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
              Altshul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
              Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
              Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
              Scapleton,M., Soares,M.B., Bonaldo,M.P., Casavant,T.L.,
              Schmetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
              Carrinci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
              Abranson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
              McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
              Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
              Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
              Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S.,
              Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
              Buffard,G.G., Blakesley,R.W., Touchman,O.W., Green,E.D.,
              Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
              Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smallu,D.E.,
              Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
TITLE        Generation and initial analysis of more than 15,000 full-length
              human and mouse cDNA sequences
JOURNAL        Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PUBMED       12477932
REFERENCE    2 (bases 1 to 1109)
AUTHORS      Strausberg,R.
TITLE        Direct Submission
JOURNAL      Submitted (02-DEC-2003) National Institutes of Health, Mammalian
              Gene Collection (MGC), Cancer Genomics Office, National Cancer
              Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
              USA
REMARK       NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT       Contact: MGC help desk
              Email: cgapbs-r@mail.nih.gov
              Tissue Procurement: John C. Marshall, M.D., Ph.D
              cDNA Library Preparation: CLONTECH Laboratories, Inc.
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
              DNA Sequencing by: Genome Sequence Centre,
              BC Cancer Agency, Vancouver, BC, Canada
              info@bcgsc.bc.ca
              Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
              Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
              Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
              Ness, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
              Ness, Pawan Pandoh, Anna-Liisa Prabh, Parvaneh Saedi, Jacqueline
              Schein, Duane Smallu, Michael Smith, Lorraine Spence, Jeff Scott,
              Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
              George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 52 Row: 6 Column: 6
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 19705452.

FEATURES             source
1..1109
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="MGC:72699 IMAGE:6922267"
/tissue_type="Pituitary gland, anterior, rat"
/clone_lib="NICHD Rr Pit1"
/lab_host="DH10B"
/notes="Vector: pDNR-LIB"
70..537
/codon_start=1
/product="Unknown (protein for MGC:72699)"
/protein_id="AAH63150.1"
/db_xref="GI:38649092"
/translation="MADKQMDNEVLMFTSYATIILAKOMFLSSATQRLTNKVF
ANPDCAGFGKGENAKFLRDEKVRRAHLNDLENIVPFLGILSLSGEDLST
ALIHFRIFVGARIYHTIATLTFLPQPNRGLAFLVGVGYGWTLSMAYLLRSRLY"
94..531
/misc_feature
/notes="MAPEG; Region: MAPEG family. This family is has
been called MAPEG (Membrane Associated Proteins in

```

Eicosanoid and Glutathione metabolism). It includes proteins such as prostaglandin H synthase. This enzyme catalyzes the synthesis of PGE2 from PGH2 (produced by cyclooxygenase from arachidonic acid). Because of structural similarities in the active sites of FLAP, LTC4 synthase and PGE synthase, substrates for each enzyme can compete with one another and modulate synthetic activity" /db_xref="CDD:pfam01124"

ORIGIN

Query Match 1.3%; Score 49; DB 10; Length 1109;
Best Local Similarity 100.0%; Pred. No. 5.9e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3714 AAATGAAAAA 1154 bp mRNA linear PRI 11-DEC-2002
Db 893 AAATGAAAAA 1154 bp mRNA linear PRI 11-DEC-2002

RESULT 43

BC040907 1154 bp mRNA linear PRI 11-DEC-2002
LOCUS BC040907 Homo sapiens, clone IMAGE:5744167, mRNA.
DEFINITION BC040907
ACCESSION BC040907
VERSION BC040907.1 GI:26454774
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM

REFERENCE

1 (bases 1 to 1154)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Straussberg, R.
Direct Submission
Submitted (06-DEC-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapsb@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www.shgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAC Plate: 88 Row: 1 Column: 1.

FEATURES

source

1. .1154
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5744167"
/tissue_type="Brain, adult medulla"
/clone_lib="NIH MGC 119"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"

ORIGIN

Query Match 1.3%; Score 49; DB 9; Length 1154;
Best Local Similarity 100.0%; Pred. No. 5.9e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3714 AAATGAAAAA 1154 bp mRNA linear PRI 11-DEC-2002
Db 1090 AAATGAAAAA 1154 bp mRNA linear PRI 11-DEC-2002

RESULT 44

BC056295

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

COMMENT

JOURNAL

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

COMMENT

JOURNAL

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

COMMENT

JOURNAL

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

COMMENT

JOURNAL

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

COMMENT

JOURNAL

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

COMMENT

JOURNAL

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

COMMENT

JOURNAL

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

COMMENT

JOURNAL

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

COMMENT

JOURNAL

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

COMMENT

BC056295 1320 bp mRNA linear VFT 08-OCT-2003
Danio rerio cDNA clone MGC:65806 IMAGE:6792135, complete cds.
BC056295
BC056295.1 GI:33604046
MGC
Danio rerio (zebrafish)
Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.

1 (bases 1 to 1320)
Straussberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, F., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.P., Casavant, T.L., Schetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., R.A., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalls, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A. M.A. full-length Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

23388257
12477932
2 (bases 1 to 1320)
Straussberg, R.
Direct Submission
Submitted (11-AUG-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapsb@mail.nih.gov
Tissue Procurement: Leonard I. Zon, M.D.
CDNA Library Preparation: Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@hgr.nih.gov
Alkhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W., Turgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAC Plate: 121 Row: e Column: 21
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis. Similarity but not identity to protein.
Location/Qualifiers
1. .1320
/organism="Danio rerio"

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAC Plate: 121 Row: e Column: 21
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis. Similarity but not identity to protein.
Location/Qualifiers
1. .1320
/organism="Danio rerio"

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAC Plate: 121 Row: e Column: 21
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis. Similarity but not identity to protein.
Location/Qualifiers
1. .1320
/organism="Danio rerio"

FEATURES

source

JOURNAL

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

COMMENT

JOURNAL

TITLE


```

FEATURES             Ipsogen (FR)
source               Location/Qualifiers
1..1429
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/notes="primer"
misc_feature         1..1429
/notes="protein phosphatase 4 (formerly x), catalytic
subunit (PPP4C) gene."
ORIGIN
Query Match          1.3%; Score 49; DB 6; Length 1429;
Best Local Similarity 100.0%; Pred. No. 5.8e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3714 AAATGAAAAA.....1429 bp mRNA linear PRI 17-DEC-1998
|||||
Db 1336 AAATGAAAAA.....1429 bp mRNA linear PRI 17-DEC-1998
|||||
RESULT 47
LOCUS AF097996 1429 bp mRNA linear PRI 17-DEC-1998
DEFINITION Homo sapiens protein phosphatase X (PPX) mRNA, complete cds.
ACCESSION AF097996
VERSION AF097996.1 GI:4028574
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1429)
AUTHORS Hu,M.C., Tang-Oxley,Q., Qiu,W.R., Wang,Y.P.,
Mihindukulasuriya,K.A., Afshar,R. and Tan,T.H.
TITLE Protein phosphatase X interacts with c-Rel and stimulates
c-Rel/nuclear factor kappaB activity
JOURNAL J. Biol. Chem. 273 (50), 33561-33565 (1998)
MEDLINE 99057922
PUBMED 9837938
REFERENCE 2 (bases 1 to 1429)
AUTHORS Hu,M.C.-T., Tang-Oxley,Q., Qiu,W., Wang,Y.-P.,
Mihindukulasuriya,K.A., Afshar,R. and Tan,T.-H.
TITLE Direct Submission
JOURNAL Submitted (13-OCT-1998) Department of Cell Biology, Amgen, Inc.,
One Amgen Center Road, 141-D, Thousand Oaks, CA 91320, USA
FEATURES             Location/Qualifiers
source               1..1429
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..1429
/gene="PPX"
105..1028
/gene="PPX"
/codon_start=1
/product="protein phosphatase X"
/protein_id="AAC96318.1"
/db_xref="GI:4028575"
/translation="MAESDLDRLQELRCRELKESVAKLCAKARELLVESNVOR
VDSPTVCGDILHGQDIFDKELFRVGVDFVDFVDRGFYSVETILLALKV
RYPDRTILRHNSRQITQYGFYDECLRYGVTVWRYCTIFDYLSAISIDGKI
FCVHGGLSPSIQTLDIITIDRKQVPHDGMCDLMSDPEDTTGWSVPRGAGYLF
SDVAQFNAANDIMICRAHQLVMEGVKHFNETVLTWSAPNYCYRCGNVAILELD
EHLQKDFIIFEAAPQETRGIPSKPVDYFL"
ORIGIN
Query Match          1.3%; Score 49; DB 9; Length 1429;
Best Local Similarity 100.0%; Pred. No. 5.8e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3714 AAATGAAAAA.....1429 bp mRNA linear PRI 17-DEC-1998
|||||

```

```

Db 1336 AAATGAAAAA.....1440 bp mRNA linear INV 01-NOV-2003
|||||
RESULT 48
LOCUS AY297438 1440 bp mRNA linear INV 01-NOV-2003
DEFINITION Aiptasia pulchella beta actin mRNA, complete cds.
ACCESSION AY297438
VERSION AY297438.1 GI:34148150
KEYWORDS Aiptasia pulchella
SOURCE Aiptasia pulchella
ORGANISM Aiptasia pulchella
REFERENCE 1 (bases 1 to 1440)
AUTHORS Lin,C., Tsai,J., Chen,M., Wang,L. and Fang,L.
TITLE Characterization of 5'-flanking region of the anemone ATP
ribosylation factor 1 and actin genes
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1440)
AUTHORS Lin,C., Tsai,J., Chen,M., Wang,L. and Fang,L.
TITLE Direct Submission
JOURNAL Submitted (13-MAY-2003) Research & Planning, National Museum of
Marine Biology & Aquarium, 2 Houwan Road, Checheng, Pintung 944,
Taiwan
FEATURES             Location/Qualifiers
source               1..1440
/organism="Aiptasia pulchella"
/mol_type="mRNA"
/db_xref="taxon:12924"
/notes="authority: Aiptasia pulchella Carlgren 1943"
74..1198
/codon_start=1
/product="beta actin"
/protein_id="AAQ62633.1"
/db_xref="GI:34148151"
/translation="MMMLPLLFDNGSGMCKDFAGDAPRAVFPISVGRPRHQGMV
GMQKDSYVGDEAQSKRGILTLKYEIHGIVTNDMDMEKIWHHTFVNSLRVAPRHPV
LATEAPLANKREKWTQIMFETNSPAMVIAQAVLSYASGRITGIVFSDGQVSH
TVPIYEGVALPHAIRLDLAGDLTDYLMKILTERGYSTTAAREIVRDILKEKLCYV
AUDFEQMOTAASSSSSELYELPGQVITIGNERFCPEALPQSFGLMESAGHET
TYSIMKCDADIRKDIYANTVLSGGSTMPFGIADRMQKEITSLAPTTWKIKLIAPPER
KYSVMIGGSILASLSTFQOMWISQKOEYDESGPSIVHRKCF"
ORIGIN
Query Match          1.3%; Score 49; DB 3; Length 1440;
Best Local Similarity 100.0%; Pred. No. 5.8e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3714 AAATGAAAAA.....1440 bp mRNA linear PRI 21-OCT-2003
|||||
Db 1388 AAATGAAAAA.....1436
|||||
RESULT 49
LOCUS BC042195 1499 bp mRNA linear PRI 21-OCT-2003
DEFINITION Homo sapiens malonyl-CoA:acyl carrier protein transacylase
(malonyl:transferrase), mRNA (cdna clone MGC:47838 IMAGE:6052380),
complete cds.
ACCESSION BC042195
VERSION BC042195.1 GI:27502988
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1499)
AUTHORS Strausberg,K.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,K.H., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,

```


Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Schetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahy, J., Helton, E., Kettman, M., Madan, A.C., Rodrigues, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smal, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

22388257

12477932

2 (bases 1 to 1499)

Strausberg, R.

Direct Submission

Submitted (02-JAN-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgaps-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland

Web site: <http://www.nisc.nih.gov/>

Contact: nisc.mgc@nci.nih.gov

Akhter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.L., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W., Tsurgueon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAC Plate: 82 Row: 0 Column: 7

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 27477044.

Location/Qualifiers

1. .1499

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:47838 IMAGE:5052380"

/tissue_type="Testis, embryonal carcinoma"

/clone_lib="NIH MGC 92"

/lab_host="DH10B"

/notes="Vector: pCMV-SPORT6"

1. .1499

/gene="MT"

/notes="synonyms: BK1191B2.3, MGC47838"

/db_xref="LocusID:27349"

214. .1215

/codon_start=1

/product="malonyl-CoA:acyl carrier protein transacylase (malonyltransferase)"

/protein_id="AAH42195.1"

/db_xref="GI:27502989"

/db_xref="LocusID:27349"

/translation="MPGQSVLLFPQGSQVVMGRGLNTPRVRELYAARVILGYD LLESLHGPQETIDRTVHCQPAIFVLAALAEKHLHLPSPFIENCAVAAAGFVGEFAA

LVPAGAEFAEGLYAVKIRAEAMQSEAVPSGMLSVLQPOSKFNFACLAREHCKS LGIENPVCVSNLPPDCRVISGHQEARLRFQKSSKFHRETRMLPVSGAFPTRLME PAVEPLTQALKAVDIKKPLVSVYSNVGHRYEHPGHIHLKLAQQLVSPVKKWQTHAI YERKKGSGPQTFEVPGRQLGAILKSCNQAWKSYSVAVDLTQTHLEVDLPQEPFR" 232. .1128

misc_feature

/note="FabD; Region: (acyl-carrier-protein) S-malonyltransferase (Lipid metabolism)"

/db_xref="CDD:COG0331"

ORIGIN

Query Match 1.3%; Score 49; DB 9; Length 1499; Best Local Similarity 100.0%; Pred. No. 5.8e-13; Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3714 AAATGAAAAA 3762

Db 1370 AAATGAAAAA 1418

RESULT 50

BC008836

LOCUS

DEFINITION

Homo sapiens transient receptor potential cation channel, subfamily C, member 4 associated protein, mRNA (cDNA clone IMAGE:3530893), partial cds.

ACCESSION

BC008836

VERSION

BC008836.2

GI:39645210

KEYWORDS

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 1595)

USDA, R.D., Collins, F.S., Wagner, L.H., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Sapich, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahy, J., Helton, E., Kettman, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smal, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 1595)

Strausberg, R.

Direct Submission

Submitted (25-MAY-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

On Dec 9, 2003 this sequence version replaced gi:14250733.

Contact: MGC help desk

Email: cgaps-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland

Web site: <http://www.nisc.nih.gov/>

Contact: nisc.mcgnhri.nih.gov
 Akhter N., Ayele K., Beckstrom-Sternberg S.M., Benjamin B.,
 Blakesley R.W., Bouffard G.G., Breen K., Brinkley C., Brooks S.,
 Dietrich N.L., Granite S., Guan X., Gupta J., Haghighi P.,
 Hansen N., Ho S.-L., Karlins E., Kwong P., Laric P., Legaspi R.,
 Maduro Q.L., Masello C., Maskeri B., Mastrian S.D., McCloskey J.C.,
 McDowell J., Pearson R., Stantropop S., Thomas P.J., Touchman J.W.,
 Turgeon C., Vogt J.L., Walker M.A., Wetherby K.D., Wiggins L.,
 Young A., Zhang L.-H. and Green E.D.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
 Series: IRAL Plate: 20 Row: a Column: 9

FEATURES
 source
 1. 1595
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:3530893"
 /tissue_type="Muscle, rhabdomyosarcoma"
 /clone_lib="NIH MGC 17"
 /lab_host="DH10B-R"
 /note="vector: pOTB7"
 <1..1595
 /gene="TRPC4AP"
 /notes="synonyms: DKFZP727M231, DKFZp586C1223, dJ756N5.2,
 TRP4AP, C20orf188, TRUSS"
 /db_xref="LocusID:26133"
 <1..794
 /gene="TRPC4AP"
 /codon_start=1
 /product="TRPC4AP protein"
 /protein_id="AAH08836.2"
 /db_xref="GI:39645211"
 /db_xref="LocusID:26133"
 /translations="QARAVSFRLGTTVSADQFLKRGLEHLYCIVDSECKSRDV
 LQSFVDLGLAEVLSECELLAYISQVPTOMSEFLRLNIHVLQENVCINTSLV
 RFENQDMKYEVLSECELLAYISQVPTOMSEFLRLNIHVLQENVCINTSLV
 ILMARKERLPLVRLRLQMEHSKKYPGFLNNFNLAPFQQHLYHKDKOSTCLN
 SSCISFSYWKETVSILLNPRQSPSALVSIEEPPYMDIDRDFTEE"

gene
 1. 1595
 /gene="TRPC4AP"
 /notes="synonyms: DKFZP727M231, DKFZp586C1223, dJ756N5.2,
 TRP4AP, C20orf188, TRUSS"
 /db_xref="LocusID:26133"
 <1..794
 /gene="TRPC4AP"
 /codon_start=1
 /product="TRPC4AP protein"
 /protein_id="AAH08836.2"
 /db_xref="GI:39645211"
 /db_xref="LocusID:26133"
 /translations="QARAVSFRLGTTVSADQFLKRGLEHLYCIVDSECKSRDV
 LQSFVDLGLAEVLSECELLAYISQVPTOMSEFLRLNIHVLQENVCINTSLV
 RFENQDMKYEVLSECELLAYISQVPTOMSEFLRLNIHVLQENVCINTSLV
 ILMARKERLPLVRLRLQMEHSKKYPGFLNNFNLAPFQQHLYHKDKOSTCLN
 SSCISFSYWKETVSILLNPRQSPSALVSIEEPPYMDIDRDFTEE"

CDS
 1. 1595
 /gene="TRPC4AP"
 /notes="synonyms: DKFZP727M231, DKFZp586C1223, dJ756N5.2,
 TRP4AP, C20orf188, TRUSS"
 /db_xref="LocusID:26133"
 <1..794
 /gene="TRPC4AP"
 /codon_start=1
 /product="TRPC4AP protein"
 /protein_id="AAH08836.2"
 /db_xref="GI:39645211"
 /db_xref="LocusID:26133"
 /translations="QARAVSFRLGTTVSADQFLKRGLEHLYCIVDSECKSRDV
 LQSFVDLGLAEVLSECELLAYISQVPTOMSEFLRLNIHVLQENVCINTSLV
 RFENQDMKYEVLSECELLAYISQVPTOMSEFLRLNIHVLQENVCINTSLV
 ILMARKERLPLVRLRLQMEHSKKYPGFLNNFNLAPFQQHLYHKDKOSTCLN
 SSCISFSYWKETVSILLNPRQSPSALVSIEEPPYMDIDRDFTEE"

ORIGIN
 Query Match 1.3%; Score 49; DB 9; Length 1595;
 Best Local Similarity 100.0%; Pred. No. 5.7e-13;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3714 AAATGAAAAA... 1655 bp mRNA linear PRI 19-NOV-2003
 Db 1531 AAATGAAAAA... 1655 bp mRNA linear PRI 19-NOV-2003

RESULT 51
 BC039725
 LOCUS Homo sapiens cdna clone MGC:47828 IMAGE:5227869, complete cds.
 DEFINITION
 ACCESSION BC039725
 VERSION BC039725.1 GI:24980796
 KEYWORDS MGC
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1655)
 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
 Klausner, R.D., Collins, P.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
 Diatchenko, L., Marais, K., Farmer, A.A., Rubin, G.M., Hong, L.,
 Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
 Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
 Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J.,
 Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J.,
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,

Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
 Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
 Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butterfield, J.S., Krzywinski, M.I., Skalska, U., Smal, D.E.,
 Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 22388257
 12477932
 2 (bases 1 to 1655)
 Direct Submission
 Strausberg, R.
 Submitted (12-NOV-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: http://mgc.nci.nih.gov
 Contact: MGC help desk
 Email: cgabbs@mail.nih.gov
 Tissue procurement: Life Technologies, Inc.
 cDNA library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: the I.M.A.G.E. Consortium (LLNL)
 Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC),
 Gaithersburg, Maryland;
 Web site: http://www.nisc.nih.gov/
 Contact: nisc.mcgnhri.nih.gov
 Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
 Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
 Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
 Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
 Maduro, Q.L., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
 McDowell, J., Pearson, R., Stantropop, S., Thomas, P.J., Touchman, J.W.,
 Turgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
 Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
 Series: IRAL Plate: 82 Row: f Column: 20
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis, GenesScan gene prediction.

FEATURES
 source
 1. 1655
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:47828 IMAGE:5227869"
 /tissue_type="Pancreas, Spleen, adult pooled"
 /clone_lib="NIH MGC_120"
 /lab_host="DH105"
 /note="Vector: pCMV-SPORT6"
 191..1162
 /codon_start=1
 /product="Unknown (protein for MGC:47828)"
 /protein_id="AAH39725.1"
 /db_xref="GI:24980797"
 /translation="MQNALAVLLAFSPASQKSNLEGRKTSVIRQTGSSAETCDLA
 EGSNGYIHWYHQEAPQRYDYSKNSKVSLESGVSGFKITTYASTNNRLILRN
 LIENFGVYCAWGPDKYKLGFGSTLVVDKQDADVPKFTIFLPSATKLLQK
 AGTYCLLEKFFPDIKIHWQKSNITLGSQEGNTKNTDTPKFSWLTVPFESLDK
 EHCIVIRHNKNKIDQIIPPIKTDTVTDPKDSYKSDANDVITMPDKONWSKDN
 DTLLQLNTSAYTYTLLLLKSVVYFAITCCLLRRTAFCNGEKS"
 284..592
 /note="Igf; Region: Immunoglobulin domain variable region
 (v) subfamily"
 /db_xref="CDD:cd00099"
 611..901
 /note="Igf; Region: Immunoglobulin domain constant region

subfamily"
/db_xref="CDD:00098"

ORIGIN

Query Match 1.3%; Score 49; DB 9; Length 1655;
Best Local Similarity 100.0%; Pred. No. 5.7e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA 1661 bp mRNA linear PRI 12-NOV-2003
DB 1596 AAATGAAAAA 1644

RESULT 52
LOCUS BC025753
DEFINITION Homo sapiens myozenin 1, mRNA (cDNA clone MGC:34233 IMAGE:520455), complete cds.
ACCESSION BC025753
VERSION BC025753.1 GI:19343990
KEYWORDS MGC.
SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1661)

REFERENCE

1 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heide, F.,
Diachenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheet, T.E., Brownstein, M.J., Uedin, T.B., Toshiyuki, S.,
Carninci, P., Frange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahy, J., Helton, E., Kettner, M., Madan, A., Young, A.C., Schrock, R.,
Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzyzanski, M.I., Skalska, U., Smalhus, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences

TITLE

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL

MEDLINE

PUBMED

2 (bases 1 to 1661)

REFERENCE

1 Strausberg, R.
Submitted (06-MAR-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK

COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc.mgc@nih.gov
Akhtar, N., Avelle, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghight, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
Maduro, O.L., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,

Tsurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Place: 49 Row: d Column: 5
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 21359948.

FEATURES

SOURCE

1..1661
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:34233 IMAGE:520455"
/tissue_type="Lung, Spleen, fetal, pooled"
/clone_lib="NIH MGC_122"
/lab_host="DH108"
/note="Vector: pCMV-SPORT6"

gene

1..1661
/gene="MYOZ1"
/note="synonyms: FATZ, CS-2, MYOZ"
/db_xref="LocusID:58529"
/db_xref="MIM:605603"
366..1265
/codon_start=1
/product="myozenin 1"
/protein_id="AAH25753.1"
/db_xref="GI:19343991"
/translation="MPLSGTPAPNKKRKKSLIMELTGGQSSGLNKGKISVPRDV
MLELSLTLNRGSKMFKLRQVKEFYIENHPDFVSDSMDFKFLPTVGQLGTAG
QGFYSKNGRGSGKAGQSGQVGSQQHLSGSGAGGTGGAGAGRGAAGTA
GVGTEGSDQAGSGKHITVITYISPERAMGVDPQOKMELGIDLLAYGAKAEIPKY
KAFNPAMPYGGYKASKMTQMPKFDLPGLLSEPLVLYNQNLNRPSFNTPTPLW
SSGEFVDVNDVIGIFLDGETEL"

CDS

366..1262
/note="Calsarcin; Region: Calcineurin-binding protein
(Calsarcin). This family consists of several mammalian
calcineurin-binding proteins. The calcium- and
calmodulin-dependent protein phosphatase calcineurin has
been implicated in the transduction of signals that
control the hypertrophy of cardiac muscle and slow fibre
gene expression in skeletal muscle. Calsarcin-1 and
calsarcin-2 are expressed in developing cardiac and
skeletal muscle during embryogenesis, but calsarcin-1 is
expressed specifically in adult cardiac and slow-twitch
skeletal muscle, whereas calsarcin-2 is restricted to fast
skeletal muscle. Calsarcins represent a novel family of
sarcomeric proteins that link calcineurin with the
contractile apparatus, thereby potentially coupling muscle
activity to calcineurin activation. Calsarcin-3, is
expressed specifically in skeletal muscle and is enriched
in fast-twitch muscle fibres. Like calsarcin-1 and
calsarcin-2, calsarcin-3 interacts with calcineurin, and
the 2-disc proteins alpha-actinin, gamma-filamin, and
telethonin"

misc_feature

366..1262
/db_xref="CDD:pfam05556"

ORIGIN

Query Match 1.3%; Score 49; DB 9; Length 1661;
Best Local Similarity 100.0%; Pred. No. 5.7e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY .3714 AAATGAAAAA 1661 bp mRNA linear PAT 06-AUG-2001
DB 1535 AAATGAAAAA 1593

RESULT 53

LOCUS

AX188231 1676 bp DNA

DEFINITION

Sequence 3926 from Patent WO0142467.

ACCESSION

AX188231

12203911
 PUBMED 2 (bases 1 to 1706)
 REFERENCE Satou, Y. and Satou, N.
 AUTHORS Direct Submission
 TITLE Submitted (04-OCT-2002) Nori Satch, Kyoto University, Department of
 JOURNAL Zoology; Sakyo-Ku, Kyoto, Kyoto 606-8502, Japan
 (E-mail: satou@asci.dan.zool.kyoto-u.ac.jp, Tel: 81-75-753-4095,
 Fax: 81-75-705-1113)
 COMMENT Ciona intestinalis cDNA Project (URL:
 http://ghost.zool.kyoto-u.ac.jp/indexr1.html).
 FEATURES Location/Qualifiers
 source 1..1706
 /organism="Ciona intestinalis"
 /mol_type="mRNA"
 /db_xref="taxon:7719"
 /clone="ciad007m12"
 ORIGIN
 Query Match 1..1706
 Best Local Similarity 100.0%; Pred. No. 5.7e-13; Length 1706;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3714 AAATGAAAAA...AAA 3762
 DB 1583 AAATGAAAAA...AAA 1631
 RESULT 56
 BC012854 1729 bp mRNA linear PRI 03-OCT-2003
 LOCUS Homo sapiens actin, beta, mRNA (cDNA clone MGC:9832 IMAGE:3863361),
 DEFINITION complete cds.
 ACCESSION BC012854
 VERSION BC012854.1 GI:15277502
 KEYWORDS MGC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1729)
 AUTHORS Strausberg R.L., Feingold E.A., Grouse, L.H., Derge, J.G.,
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
 Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
 Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
 Carninci, P., Prange, C., Raha, S., Lequellano, N.A., Peters, G.J.,
 Abramson, R.D., Mullan, S.J., Bosak, S.A., McEwan, P.J.,
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
 Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
 Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y.,
 Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butlerfield, Y.S., Krzywinski, M.I., Skalska, J., Smal, D.E.,
 Scherch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 JOURNAL MEDLINE 22388257
 PUBMED 12477932
 REFERENCE 2 (bases 1 to 1729)
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (20-AUG-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: http://mgc.nci.nih.gov
 REMARK Contact: MGC help desk
 COMMENT Email: cgabbs@mail.nih.gov

Tissue Procurement: DCTD/DTP
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome
 Sequencing Center
 Center code: BCM-HGSC
 Web site: http://www.hgsc.bcm.tmc.edu/cdna/
 Contact: ang@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulsegod, H.,
 Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
 A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov
 Series: IRAC Plate: 21 Row: h Column: 1
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis.

FEATURES Location/Qualifiers
 source 1..1729
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:9832 IMAGE:3863361"
 /tissue_type="Ovary, adenocarcinoma"
 /clone_lib="NIH MGC_66"
 /lab_host="DH10B"
 /notes="Vector: pCMV-SPORT6"
 Gene 1..1729
 /genes="ACTB"
 /db_xref="LocusID:60"
 /db_xref="MIM:102630"
 CDS 3..1085
 /codon_start=1
 /product="ACTB protein"
 /protein_id="AAH12854.1"
 /db_xref="GI:15277503"
 /db_xref="LocusID:60"
 /translation="MCKAGFAGDADPRAVFPISVGRPHQVYMGQKDSYVGDEAQ
 SKGILTKYPIEHGIVTNDMEKIWHHTFYNELRVAPEHPVLITLAPLNKANLE
 KWTQIMFTFNTAMVLAQVLSVYAGRTTGI VMDSGDVTHVTPIYEGVALPHAI
 LRDLAGRLDTYKILTERGVSTPTAERIVRDIKELCVALDPEQMATRASS
 SSLEKSEYLPDQGVITIGNERPCPEALPQPSFLGMSGCIHETTFNSIMKCDVDIRK
 DLTFAVLSGGTITMFIADRMQKEITALPSTMKIKIIPPERKYSVWIGSILASL
 STFQQWMSKGYDESGPSIVHRKCF"
 misc_feature 3..1082
 /notes="actin; Region: Actin"
 /db_xref="CDD:pfam00022"
 ORIGIN
 Query Match 1..3%; Score 49; DB 9; Length 1729;
 Best Local Similarity 100.0%; Pred. No. 5.7e-13;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3714 AAATGAAAAA...AAA 3762
 DB 1680 AAATGAAAAA...AAA 1728

RESULT 57
 BC061980 1737 bp mRNA linear ROD 11-DEC-2003
 LOCUS Rattus norvegicus cDNA clone MGC:72261 IMAGE:5597975, complete cds.
 DEFINITION Rattus norvegicus (Norway rat)
 ACCESSION BC061980
 VERSION BC061980.1 GI:38541365
 KEYWORDS MGC.
 SOURCE Rattus norvegicus
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 1737)
 AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,

Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Datchenko,L., Marusina,K., Farmer,A.A., Rubin,G.W., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Schetz,T.E., Brownstein,M.J., Uedin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Itoquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gurnatne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J.J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalhus,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 1737)

Strausberg,R.

Direct Submission

Submitted (10-NOV-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgapbs@mail.nih.gov

Tissue Procurement: Jeff Green/Patru Kondaiah, NCI.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)

DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland;

Web site: <http://www.nisc.nih.gov/>

Contact: nisc.mc@hghri.nih.gov

Akter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S., Dietrich,N.L., Grantte,S., Guan,X., Gupta,J., Haghghi,P., Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R., Maduro,Q.L., Masello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C., McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W., Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>

Series: IRAK Plate: 137 Row: b Column: 17

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 31377483.

FEATURES

source

1..1737

/organism="Rattus norvegicus"

/mol_type="mRNA"

/db_xref="taxon:10116"

/clone="MGC:72261 IMAGE:5597975"

/tissue_type="Prostate, pool of NCI_CGAP_Pr30, 40, 41 and NCI_CGAP_Pr29, 39, 42"

/clone_lib="NCI CGAP_Pr49"

/lab_host="DH10B"

/note="Vector: PCMV-SPORT6.1"

46..828

/codon_start=1

/product="Unknown (protein for MGC:72261)"

/protein_id="AAH61980.1"

/db_xref="GI:39541366"

/translation="MAKESGASHNGPHEHLYPIAKDQNSPIELHTKDIRDPSL QPWSYIDPSAKTILNGKTKRVDFDIFRSMRLRGGLSGPYLRDFLHWSSDD HGEHVTGVKVAEHLVHWNPKYNTFGALKQPDIAVAFILKIGREKGEFQILL DALDKIKTKGEAPFNHFDPSCLFFACRDYWTYHGSFTTPPCBECIIVLLLKBPMTVS

CDS

misc_feature

SDQMAKLRSLFASAEENPEPVLGVNWRPPQPIKGRVVRASF*

58..819

/note="carb_anhydriase; Region: Eukaryotic-type carbonic anhydriase"

/db_xref="CDD:pfam00194"

ORIGIN

Query Match 1.3%; Score 49; DB 10; Length 1737;

Best Local Similarity 100.0%; Pred. No. 5.7e-13;

Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762

Db 1683 AAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1731

RESULT 58

AB070107

LOCUS

AB070107 1786 bp mRNA linear PRI 16-AUG-2001

Macaca fascicularis testis cDNA clone:QtsA-14034, full insert sequence.

ACCESSION

AB070107

VERSION

AB070107.1 GI:15208054

KEYWORDS

oligo capping; fis (full insert sequence).

SOURCE

Macaca fascicularis (crab-eating macaque)

ORGANISM

Macaca fascicularis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae; Cercopitheinae; Macaca.

REFERENCE

1

Kashimoto,K., Osada,N., Hida,M., Kusuda,J., Tanuma,R., Hirai,M., Terao,K. and Sugano,S.

Isolation of novel full-length cDNA clones from macaque testis cDNA libraries

Unpublished

2 (bases 1 to 1786)

Kashimoto,K., Osada,N., Hida,M., Kusuda,J. and Sugano,S.

Direct Submission

Submitted (14-AUG-2001) Katsuyuki Hashimoto, National Institute of Infectious Diseases, Division of Genetic Resources; 23-1, Toyama 1-chome, Shinjuku-ku, Tokyo 162-8640, Japan

(E-mail:khashi@nih.go.jp, URL:<http://www.nih.go.jp/yoken/genebank/>, Tel:81-3-5285-1111(ex:2120), Fax:81-3-5285-1181)

Lab host: TOP10

Vector: pME18S-FL3 (Acc.No. AB009864)

R. Site1: DraIII (CACTGTGTG)

R. Site2: DraIII (CACTGTGTG)

Description: 1st strand cDNA was primed with an oligo(dT) primer (ATGTGGCTTTTCTTTTCTTTT), double-stranded cDNA was synthesized using specific 5' and 3' primers and amplified by PCR. The PCR product was digested with SfiI and size selection was performed to exclude fragments <1.5kb. The SfiI-digested PCR product was cloned into distinct DraIII sites of pME18S-FL3. XhoI sites just outside the DraIII sites can be used to isolate the cDNA insert. Libraries were constructed by oligo-capping method (Sugano et al., University of Tokyo, Institute of Medical Sciences).

Custom primer used for sequencing

(5' end primer (CTTCTGCTCTAAAGCTCG);

3' end primer (CGACCTGCAGCTCGACAC)).

Location/Qualifiers

1..1786

/organism="Macaca fascicularis"

/mol_type="mRNA"

/db_xref="taxon:9541"

/clone="QtsA-14034"

/sex="male"

/tissue_type="testis"

/clone_lib="macaque testis cDNA library QtsA"

/dev_stage="adult"

895..1164

/codon_start=1

/product="hypothetical protein"

/protein_id="BAB63052.1"

FEATURES

source

1..1786

/organism="Macaca fascicularis"

/mol_type="mRNA"

/db_xref="taxon:9541"

/clone="QtsA-14034"

/sex="male"

/tissue_type="testis"

/clone_lib="macaque testis cDNA library QtsA"

/dev_stage="adult"

895..1164

/codon_start=1

/product="hypothetical protein"

/protein_id="BAB63052.1"

CDS


```

/db_xref="GI:15208055"
/translation="MPCALLETITDGAPEIIATIQVFMQCFLEALGKKNKOLRFALKTY
FPYTSPLAWLLQDPQDVRGRWLQTLKHSISELLREAVDQTHG"

ORIGIN
Query Match      1.3%; Score 49; DB 9; Length 1786;
Best Local Similarity 100.0%; Pred. No. 5.7e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA... 3762
D 1643 AAATGAAAAA... 1691

RESULT 59
AY118452 1806 bp mRNA linear INV 15-JUN-2002
Drosophila melanogaster AT09813 full insert cDNA.
ACCESSION AY118452
VERSION AY118452.1 GI:21428321
KEYWORDS
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
REFERENCE 1 (bases 1 to 1806)
AUTHORS Stapleton,M., Brokstein,P., Hong,L., Agbayani,A., Carlson,J.,
Champe,M., Chavez,C., Dorsett,V., Dresnek,D., Farfan,D., Frise,E.,
George,R., Gonzalez,M., Guarin,H., Krommiller,B., Li,P., Liao,G.,
Miranda,A., Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,
Patel,S., Phouanavong,S., Wan,K., Yu,C., Lewis,S.E., Rubin,G.M.
and Celniker,S.
DIRECT SUBMISSION
Submitted (07-JUN-2002) Berkeley Drosophila Genome Project,
Lawrence Berkeley National Laboratory, One Cyclotron Road,
Berkeley, CA 94720, USA
Sequence submitted by:
Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory
Berkeley, CA 94720
This clone was sequenced as part of a high-throughput process to
sequence clones from Drosophila Gene Collection 1 (Rubin et al.,
Science 2000). The sequence has been subjected to integrity checks
for sequence accuracy, presence of a polyA tail and contiguity
within 100 kb in the genome. Thus we believe the sequence to
reflect accurately this particular cDNA clone. However, there are
artifacts associated with the generation of cDNA clones that may
have not been detected in our initial analyses such as internal
priming, priming from contaminating genomic DNA, retained introns
due to reverse transcription of unspliced precursor RNAs, and
reverse transcriptase errors that result in single base changes.
For further information about this sequence, including its location
and relationship to other sequences, please visit our Web site
(http://fruitfly.berkeley.edu) or send email to
cdna@fruitfly.berkeley.edu
FEATURES
source location/Qualifiers
1..1806
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
1..1806
/gene="CG7804"
/db_xref="FLYBASE:FBgn0036496"
341..1297
/gene="CG7804"
/feature="Longest ORF"
/codon_start=1
/product="AT09813p"
/protein_id="AA049821.1"
/db_xref="GI:21428322"
/db_xref="FLYBASE:FBgn0036496"
/translation="MVFVHVSEKSGDEPIELPABEDGTLILLSTLQAQFPSSSGLOVYN

```

```

VDTKAVRGVNRGRLKSPSEETGWGEYHFCVFPKKKNROSENLENSTAKTRTEA
HLRCDLILVNTTQDRLREYFYTDVQVKAIEIKDTRSGHSGKGFVRFQSYDV
QWHLKSHSIDGRCVCEVGPVPSMGNGEPKGVFVGRCTEDIEADDRYFSGFEV
IDVFIPKDFRAFVSFVFLDYPVPRVVCGRHIIKGSVHVSTAKNKVQNKQLPQTN
NTNLDNNFKVQPNANFRMHPPANNFMSHSPHNGYQMGVNVN"

ORIGIN
Query Match      1.3%; Score 49; DB 3; Length 1806;
Best Local Similarity 100.0%; Pred. No. 5.7e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA... 3762
D 1758 AAATGAAAAA... 1806

RESULT 60
BC029874 1835 bp mRNA linear PRI 07-OCT-2003
Homo sapiens protein inhibitor of activated STAT protein PIAS1,
mRNA (cDNA clone MGC:35296 IMAGE:5176540), complete cds.
ACCESSION BC029874
VERSION BC029874.1 GI:20987516
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1835)
AUTHORS Strausberg,R.D., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Schett,T.B., Brownstein,M.J., Uesdin,T.B., Toshiyuki,S.,
Carrincci,P., Frange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullaly,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shcherenko,Y.,
Boutard,G.G., Blakeley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smillius,D.E.,
Scherer,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
PROC. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22386257
JOURNAL MEDLINE
PUBMED 12477932
REFERENCE 2 (bases 1 to 1835)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL
Submitted (06-MAY-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulsegad, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.

```


Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 51 Rows: m Column: 22.

FEATURES

source

Location/Qualifiers: 22.
1..1835
/organism="Homo sapiens"
/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:35296 IMAGE:5176540"

/tissue_type="Brain, Lung, Testis, adult, pooled whole"

/clone_lib="NIH MGC_115"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

gene

1..1835

/gene="FIASy"

/note="synonym: MGC35296"

/db_xref="LocusID:51588"

/db_xref="MIM:605989"

11..1543

CDS

/product="protein inhibitor of activated STAT protein

FIASy"

/protein_id="AAH29874.1"

/db_xref="GI:20987517"

/db_xref="LocusID:51588"

/translation="MAELVAKNMVMSFRVSDLOMLGFVGRSGKGLKHELVTRALQ

LVQFDSPELFKKIKELVETRYAKKNSEPAQHPRLDPLTMHSTYDRAGAVPTPLA

GNIDYVLYGKXGLGRLPAKTLKPEVRLVQLPFENMLDELLKPTLVPPNNKQLQ

ESFCLIPATPROVELIRNSRELQPGKAVQVRLICYSIDTSCPDQDPYFNIAVKNH

SYCSVPGIYFNSKNGVPERPCRPINLTHLMTSATNRKTIWNGVYKSVLYLV

RLQTSSELQRLKTIQVHPELKAIVKELRLDPDSIATTVGRVSLICPLVKNRLS

VFCRAETCAHLCQCPDAPVFLQWNEKKPTWPCVCDKAPYQDLIIDGLLSLSECD

ADEIVLVGDSGCPPIRAEKERSCPGAILVLGPDANGLLPAPSVNGSGALGSGTGG

GPVSGMNGKPGADVLDLTLDSSSSSEDEEEDEEGEPKPKRCFPQKGLVP

AC"

misc_feature

44..148

/note="SAP; Region: Putative DNA-binding (bhelical) motif

predicted to be involved in chromosomal organisation"

/db_xref="CDD:smart00513"

974..1132

misc_feature

/note="zf-MIZ; Region: pfam02891, zf-MIZ, MIZ zinc finger"

/db_xref="CDD:pfam02891"

ORIGIN

Query Match 1.3%; Score 49; DB 9; Length 1835;

Best Local Similarity 100.0%; Pred. No. 5.6e-13;

Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAA 3762

DB 1732 AAATGAAA 1780

RESULT 61

BC061320

LOCUS

BC061320 1838 bp mRNA linear VRT 04-NOV-2003

Silurana tropicalis cDNA clone MGC:75804 IMAGE:5382264, complete

DEFINITION

cds.

ACCESSION

BC061320

VERSION

MGC.

SOURCE

ORGANISM

Silurana tropicalis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;

Xenopodinae; Silurana.

REFERENCE

1 (bases 1 to 1838)

AUTHORS

Klein, S.L., Strausberg, R.L., Wagner, L., Pontius, J., Clifton, S.W.

and Richardson, P.

Genetic and genomic tools for Xenopus research: The NIH Xenopus

initiative

JOURNAL

Dev. Dyn. 225 (4), 384-391 (2002)

MEDLINE

PUBMED

REFERENCE

AUTHORS

Location/Qualifiers

2 (bases 1 to 1838)

Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,

Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,

Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,

Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,

Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,

Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,

Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,

Abramson, R.D., Mullaby, S.J., Bosak, S.A., McEwan, P.J.,

McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,

Morley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,

Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,

Fahey, J., Helton, E., Kettman, M., Madan, A., Young, A.C., Shevchenko, Y.,

Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,

Bouffard, G.C., Blakesley, R.W., Touchman, J.W., Green, E.D.,

Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,

Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smallos, D.E.,

Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length

human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

22388257

3 (bases 1 to 1838)

Strausberg, R.

Direct Submission

Submitted (03-NOV-2003) National Institutes of Health, Mammalian

Gene Collection (MGC) Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: gcapsb@mail.nih.gov

Tissue Procurement: Robert M. Grainger

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing: The I.M.A.G.E. Consortium (LLNL)

Center, Stanford University School of Medicine, Stanford, CA 94305

Web site: <http://www-shgc.stanford.edu>Contact: (Dickson, Mark) mdcpaxil@stanford.edu

Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,

R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAK Plate: 143 Row: g Column: 7

This clone was selected for full length sequencing because it

passed the following selection criteria: Hexamer frequency ORF

analysis. Similarity but not identity to protein.

Location/Qualifiers

1..1838

/organism="Silurana tropicalis"

/mol_type="mRNA"

/db_xref="taxon:8364"

/clones="MGC:75804 IMAGE:5382264"

/tissue_type="Embryo, Silurana tropicalis, neurula (stages

14-19)"

/clone_lib="NICHD_XGC_Embc"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6.1"

128..1000

/codon_start=1

/product="Unknown (protein for MGC:75804)"

/protein_id="AAH61320.1"

/db_xref="GI:38174056"

/translation="MSTDGKGRKRYKSCGNKQKRRHLEVGQGLITCNWNE

RKCAEAYSLNIEYQMGYPEKLSKDGSEEDDDAEALKEVDQITSTEK

NLRFPQVSGANNVPIRTLVNPEPKLVHILKDVHTTKTKTIRLMPLVSGTCK

AFLEDKLYAETFPAPWPKFNKGTQIAYKAENNNHNRDEVIKELAGIVASQNPEN

KVDLSNPEYTVIVEIIKNVCCFSVWDYTVFRKYNLQEVTKSKSEKPPQOPTKEKEE

128..1000

/product="Unknown (protein for MGC:75804)"

/protein_id="AAH61320.1"

/db_xref="GI:38174056"

/translation="MSTDGKGRKRYKSCGNKQKRRHLEVGQGLITCNWNE

RKCAEAYSLNIEYQMGYPEKLSKDGSEEDDDAEALKEVDQITSTEK

NLRFPQVSGANNVPIRTLVNPEPKLVHILKDVHTTKTKTIRLMPLVSGTCK

AFLEDKLYAETFPAPWPKFNKGTQIAYKAENNNHNRDEVIKELAGIVASQNPEN

KVDLSNPEYTVIVEIIKNVCCFSVWDYTVFRKYNLQEVTKSKSEKPPQOPTKEKEE

```

misc_feature
NLKETQGTQNTAHS
578..859
/note="THUMP; Region: THUMP domain. The THUMP domain is
named after thioluridine synthases, methylases and
PSUs. The THUMP domain consists of about 110 amino acid
residues. It is predicted that this domain is an
RNA-binding domain that adopts an alpha/beta fold similar
to that found in the C-terminal domain of translation
initiation factor 3 and ribosomal protein S8. The THUMP
domain probably functions by delivering a variety of RNA
modification enzymes to their targets"
/db_xref="CDD:pfam02926"

ORIGIN
Query Match 1.3%; Score 49; DB 5; Length 1838;
Best Local Similarity 100.0%; Pred. No. 5.6e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AATGTAAGTAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
DB 1778 AATGTAAGTAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1826

RESULT 62
BC044938 1850 bp mRNA linear PRI 01-DEC-2003
DEFINITION Homo sapiens cDNA clone IMAGE:5190011, complete cds.
ACCESSION BC044938
VERSION BC044938.1 GI:27924313
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 1850)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, P.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.P., Zesberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldi, M.P., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raja, S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullen, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettaman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalls, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 1850)
Strausberg, R.
Direct Submission
Submitted (23-JAN-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cagps-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca

```

Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Letticia Haio, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice Moleavy, Steven Ness, Pawan Pandon, Anna-Liisa Prabhu, Farvaneh Saeedi, Jacqueline Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Teai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov> Series: IRAX Plate: 87 Row: j Column: 14 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 1630585 This clone has the following problem: The cds is short compared to the longest cds in the locus.

```

FEATURES
Location/Qualifiers
1..1850
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5190011"
/tissue type="Colon, Kidney, Stomach, adult, whole pooled"
/clone_lib="NIH MGC_116"
/lab_host="DH10B"
/notes="Vector: PCMV-SPORT6"
371..583
/codon_start=1
/product="Unknown (protein for IMAGE:5190011)"
/protein_id="AAH4938.1"
/db_xref="GI:27924314"
/translation="MNVSSSHFVSALTVPFNGSKLSIKIEDTPVDPDSLKILIVANNS
DTLLKMSCHPVSSDDGLHFLKLE"
ORIGIN
Query Match 1.3%; Score 49; DB 9; Length 1850;
Best Local Similarity 100.0%; Pred. No. 5.6e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AATGTAAGTAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
DB 1735 AATGTAAGTAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1783

RESULT 63
AB063046 1894 bp mRNA linear PRI 13-JUN-2001
LOCUS Macaca fascicularis brain cDNA clone:QMOA-11972, full insert
sequence.
ACCESSION AB063046.1 GI:14388491
VERSION AB063046
KEYWORDS oligo capping; fis (full insert sequence)
SOURCE Macaca fascicularis (crab-eating macaque)
ORGANISM Macaca fascicularis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
Cercopithecinae; Macaca.
REFERENCE
1 (sites)
Osada, N., Hida, M., Kusuda, J., Tanuma, R., Iseki, K., Hirai, M.,
Terao, K., Suzuki, Y., Sugano, S. and Hashimoto, K.
Isolation of full-length cDNA clones from macaque brain cDNA
libraries
Unpublished
2 (bases 1 to 1894)
Hashimoto, K., Osada, N., Hida, M., Kusuda, J. and Sugano, S.
Direct Submission
Submitted (11-JUN-2001) Katsuyuki Hashimoto, National Institute of
Infectious Diseases, Division of Genetic Resources; 23-1, Toyama
1-chome, Shinjuku-ku, Tokyo 162-8640, Japan
(E-mail: khashi@nih.go.jp, URL: http://www.nih.go.jp/yoken/genebank/,
Tel: 81-3-5285-1111 (ex.2120), Fax: 81-3-5285-1181)
Lab host: TOP10
Vector: pME18S-FL3 (Acc.No. AB009864)
COMMENT

```

R. Site1: DraIII (CACTGTGTG)
 R. Site2: DraIII (CACCATGTG)
 Description: 1st strand cDNA was primed with an oligo(dT) primer
 [ATGCGCCCTTTTCTTTTCTTTT]; double-stranded cDNA was synthesized
 using specific 5' and 3' primers and amplified by PCR. The PCR
 product was digested with SfiI and size selection was performed to
 exclude fragments <1.5kb. The SfiI-digested PCR product was cloned
 into distinct DraIII sites of pME18S-FL3. XhoI sites just outside
 the DraIII sites can be used to isolate the cDNA insert. Libraries
 were constructed by oligo-capping method
 (Sugano et al., , Institute of Medical Science, University of
 Tokyo).

Custom primer used for sequencing
 (5' end primer [CTTCTGCTCTAAAGCTGCG];
 3' end primer [CGACTGCTGAGCTGAGCACA]).

FEATURES

source
 Location/Qualifiers
 1..1894
 /organism="Macaca fascicularis"
 /mol_type="mRNA"
 /db_xref="taxon:9541"
 /clone="QmoA-11972"
 /sex="male"
 /tissue_type="medulla oblongata"
 /clone_lib="macaque brain cDNA library QmoA"
 /dev_stage="adult"

ORIGIN

Query Match 1.3%; Score 49; DB 9; Length 1894;
 Best Local Similarity 100.0%; Pred. No. 5.6e-13;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAA 3762
 DB 1773 AAATGAAA 1821

RESULT 64

AR300635
 LOCUS AR300635 1933 bp DNA linear PAT 12-JUN-2003
 DEFINITION Sequence 3 from patent US 6537811.
 ACCESSION AR300635
 VERSION AR300635.1 GI:31688184
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.

REFERENCE

1 (bases 1 to 1933)
 Freier,S.M.
 TITLE Antisense inhibition of SAP-1 expression
 JOURNAL Patent: US 6537811-A 3 25-MAR-2003;
 LOCATION/Qualifiers

FEATURES

source
 1..1933
 /organism="unknown"
 /mol_type="genomic DNA"

ORIGIN

Query Match 1.3%; Score 49; DB 6; Length 1933;
 Best Local Similarity 100.0%; Pred. No. 5.6e-13;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAA 3762
 DB 1878 AAATGAAA 1926

RESULT 65

AX146886
 LOCUS AX146886 1933 bp DNA linear PAT 08-JUN-2001
 DEFINITION Sequence 20 from Patent WO0136674.
 ACCESSION AX146886
 VERSION AX146886.1 GI:14346186
 KEYWORDS
 SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1
 Bull,J.H., Ellison,G. and Paskins,L.D.
 TITLE Diagnostic methods for the detection of prostate disorders
 JOURNAL Patent: WO 0136674-A 20 25-MAY-2001;
 Astrazeneca AB (SE)

FEATURES

Location/Qualifiers
 1..1933
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN

Query Match 1.3%; Score 49; DB 6; Length 1933;
 Best Local Similarity 100.0%; Pred. No. 5.6e-13;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAA 3762
 DB 1878 AAATGAAA 1926

RESULT 66

HUMSAPIA
 LOCUS HUMSAPIA 1933 bp mRNA linear PRI 17-DEC-1993
 DEFINITION Homo sapiens SRP accessory protein 1A (SAP-1) mRNA, complete cds.
 ACCESSION M85165
 VERSION M85165.1 GI:429185
 KEYWORDS serum response factor; SAP-1; Elk-1.
 SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 1933)
 Dalton,S. and Treisman,R.
 TITLE Characterization of SAP-1, a protein recruited by serum response
 factor to the c-fos serum response element
 JOURNAL Cell 68 (3), 597-612 (1992)
 MEDLINE 92154673
 PUBMED 1339307

REFERENCE

2 (bases 1 to 1933)
 Treisman,R.
 TITLE Direct Submission

AUTHORS

Imperial Cancer Research Fund, London, England
 Submitted (05-MAR-1992) Richard Treisman, Transcription Laboratory,
 On Nov 29, 1993 this sequence version replaced gi:338036.

COMMENT

Original source text: Homo sapiens cDNA to mRNA.

FEATURES

Location/Qualifiers
 1..1933
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"

gene

1..1933
 /gene="SAP-1"

CDS

150..1445
 /gene="SAP-1"
 /note="Homology region A with Elk-1 protein (Ets domain)
 is bp 150-417, amino acids 1-89; Homology region B with
 Elk-1 protein, required for cooperative ternary complex
 formation with SRF is bp 565-617, amino acids 136-157;
 Sequence diverges from SAP-1B at bp 1229, amino acid 360;
 Homology region C with Elk-1 protein, core of regulated
 transcription activation domain, is bp 1203-1355, amino
 acids 352-402; (S/T)P motifs conserved between SAP-1A and
 Elk-1 are located at amino acids T354, T361, T366, S381,
 S387, T420, S425, corresponding to bp 1209-1214, bp
 1230-1235, bp 1245-1250, bp 1290-1295, bp 1308-1313, bp
 1407-1412, bp-1422-1427"
 /codon_start=1
 /product="SAP-1A protein"
 /protein_id="AAA03631.1"


```
source
1. 2015
/organism="Aedes aegypti"
/mol_type="mRNA"
/strain="Rockefeller"
/db_xref="taxon:7159"
/tissue_type="whole body"
/dev_stages="larva"
4. 1905
/function="iron-binding protein"
/function="iron-transport protein"
/codon_start=1
/product="transferrin precursor"
/protein_id="AAB87414.1"
/db_xref="GI:2645497"
/translation="MTVSNALLSVVAVILLIGSHGQOKETFKLPHQIMDACQDLM
AKPDAIOVCIAGRDMECKVKAREADFAVDPMVTVAYHMANODSVTEFTF
LEPKAEPRVEGIIILVRKDNFSLADLKGKSGHTGVRNGVYKIPITKLKSGAFK
LATSELSPEKELKGLSLDFGSACLVGKSPNDEVNRLKKRYSNLCALCEPVED
YDKSGTGDGALRCLVNNGDVATKIYVKNYFGLFVGGAAPALNPARTEDYV
ICEGSTRPITGPACSWAQRWQGTMGNDINSRFLQRLQOQFYODAKNSADTKA
LKMVDRKNVLVDREVVPQPGDHLNRAQYDVIERDFQNKIKLVTSLIENKCEV
MKAAYSRDVPAFCVCMKGSCVEAVERGEADVVLKGEDQPATSTSELKILFEK
YDDNMVAVADKQITRDQVLKAPLEFADSPRAVSAALFNDKQKQACPNKLASQ
GAPRIVGAQDLTRIGANKRLICPNLEQALGNVQACNVDTYMTPTGIYVRKDTGQLE
DNIAHAFVLSDKZGHGAKNEVVEFGEFAGEKNVLFHDBRAKFGVAGNKIGEDVP
ANIRKLMCL"
sig_peptide
mat_peptide
67. 1902
/product="transferrin"

ORIGIN
Query Match 1.3%; Score 49; DB 3; Length 2015;
Best Local Similarity 100.0%; Pred. No. 5.6e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA... 3762
Db 1966 AAATGAAAAA... 2014

RESULT 71
BC053219 2031 bp mRNA linear VRT 08-OCT-2003
LOCUS
DEFINITION
Danio rerio cDNA clone MGC:64043 IMAGE:6793489, complete cds.
ACCESSION
BC053219.1 GI:31419505
VERSION
MGC:
KEYWORDS
Danio rerio (zebrafish)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.

REFERENCE
1 (bases 1 to 2031)
Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan K., Moore T., Max S.I., Wang J., Hsieh F.,
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L.,
Schaeetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S.,
Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J.,
Abramson R.D., Mullaney S.J., Bosak S.A., McEwan P.J.,
McKernan K.J., Malek J.A., Gunaratne P.H., Richards S.,
Worley K.C., Hale S., Garcia A.M., Gay L.J., Rulyk S.W.,
Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Fahey J., Helton E., Kettman M., Madan A.C., Rodrigues S.,
Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y.,
Bouffard G.G., Blakeley R.W., Touchman J.W., Green E.D.,
Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
Butterfield Y.S., Krzywinski M.I., Skalska U., Smallus D.E.,
Scherch A., Schein J.E., Jones S.J. and Marra M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
```

```
JOURNAL
MEDLINE
22388257
PUBMED
12477932
REFERENCE
2 (bases 1 to 2031)
Strausberg R.
Direct Submission
Submitted (02-JUN-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapps-remail.nih.gov
Tissue Procurement: Leonard I. Zon, M.D.
cDNA Library Preparation: Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www.shgc.stanford.edu
Contact: (Dickson, Mark) mcdpaxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LINL at: http://image.llnl.gov
Series: IRAK Plate: 117 Row: i Column: 8
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, Similarity but not identity to protein.
FEATURES
Location/Qualifiers
1..2031
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="MGC:64043 IMAGE:6793489"
/tissue_type="Kidney, zebrafish"
/clone_lib="NCI CGAP_ZKId1"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6.1"
140..514
/codon_start=1
/product="Unknown (protein for MGC:64043)"
/protein_id="AAH53219.1"
/db_xref="GI:31419506"
/translaton="MAKTDGSSFFRRSSVFLMTVSLALAFFTWTFVPRDYPYGSGLG
PLGALANYVVEOHYTVYVYGFETLHLCEAMFALKLCSKDGIFSSSARLLWFVQTF
LFGFASLGLLLKYPDGRKHH"

CDS
Query Match 1.3%; Score 49; DB 5; Length 2031;
Best Local Similarity 100.0%; Pred. No. 5.6e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA... 3762
Db 1948 AAATGAAAAA... 1996

RESULT 72
HSU55017 2038 bp mRNA linear PRI 05-MAY-1996
LOCUS
DEFINITION
Human transketolase (TKT) mRNA, complete cds.
ACCESSION
U55017 M86521
VERSION
U55017.1 GI:1297296
KEYWORDS
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2038)
Abedinia M., Layfield R., Jones S.M., Nixon P.F. and Mattick J.S.
Nucleotide and predicted amino acid sequence of a cDNA clone
encoding part of human transketolase
```


JOURNAL Biochem. Biophys. Res. Commun. 183 (3), 1159-1166 (1992)
 MEDLINE 92231878
 PUBMED 1567394
 REFERENCE 2 (bases 1 to 2038)
 AUTHORS Schenk, G., Layfield, R., Candy, J.M., Duggleby, R.G. and Nixon, P.F.
 TITLE Molecular evolutionary analysis of the thiamin-dependent enzyme, transketolase
 JOURNAL Unpublished
 REFERENCE 3 (bases 1 to 2038)
 AUTHORS Schenk, G., Candy, J.M., Duggleby, R.G. and Nixon, P.F.
 TITLE Direct Submission
 JOURNAL Submitted (26-MAR-1996) G. Schenk, Biochemistry, University of Queensland, Brisbane, Qld 4072, Australia
 COMMENT On May 6, 1996 this sequence version replaced gi:339941.
 FEATURES Location/Qualifiers
 source
 1. .2038
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /chromosome="3"
 /map="3p14.3"
 /tissue_type="brain"
 1. .2038
 /gene="TKT"
 1. .1872
 /gene="TKT"
 /EC_number="2.2.1.1"
 /codon_start=1
 /product="transketolase"
 /protein_id="AAA98961.1"
 /db_xref="GI:1297297"
 /translation="MESYHKPDQKQLQALKDTANLRISSIQTAAGSGHPTSCCSA
 AEINAVLFFHTMYKSDQPNRPNDRFVLSKGHAAPILYAVMAEAGFLAEALLNRK
 ISSLDGHPVPKQAFDTATGSLGGLGACGMAYTKYFDKASRYVYCLLDGGLSE
 GSVWEAMAFASYIKLDNLVAIDLINRLGSDPAPLOHMDIYQKCEAFGWHAIIYDG
 HSEVELKAFQAKHQPTAIATKTFKRGITGVDEKESWGKPLPKMAEQIIQBIYS
 IQQSKKLTAPDAPSDVDIANIRMPSLPSYKVGDKIATRAYGQALAKLGHASDR
 IIALDGTNTSTFSEIFKKEHPDRFIECYIAEQNNVSIAGVCATRNTPVFCSTFAAF
 FTRAPDQIRMAAISININLCSHGVSIGEDGSPQMALEDLAFERSVPTSTVFPSD
 GVTEKAVELAAANTKGICFIRTPENAIYNNEDFQVQAKVLSKDDQVTVIGA
 GVTLEHAAALAEKLLKKNIRVLPDFTKPLDKLILDSARATKGRILTVEDHYVIG
 GIGEAIVSSAVGEPGIVTTHLVANVRPSRGPPELLKMFIDRDAIAQAVRGLITKA"
 ORIGIN
 Query Match 1.3%; Score 49; DB 9; Length 2038;
 Best Local Similarity 100.0%; Pred. No. 5.6e-13;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3714 AAATGAAAAA... 3762
 Db 1974 AAATGAAAAA... 2022
 RESULT 73
 HSTRANSK
 LOCUS HSTRANSK 2106 bp mRNA linear PRI 09-MAY-1996
 DEFINITION H.sapiens mRNA for transketolase.
 ACCESSION X67688 S52775
 VERSION X67688.1 GI:37266
 KEYWORDS transketolase
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 2104)
 AUTHORS Abedin, M., Layfield, R., Jones, S.M., Nixon, P.F. and Mattick, J.S.
 TITLE Nucleotide and predicted amino acid sequence of a cDNA clone encoding part of human transketolase
 JOURNAL Biochem. Biophys. Res. Commun. 183 (3), 1159-1166 (1992)
 MEDLINE 92231878
 PUBMED 1567394
 REFERENCE 2
 AUTHORS McCool, B.A., Plonk, S.G., Martin, P.R. and Singleton, C.K.

TITLE Cloning of human transketolase cDNAs and comparison of the nucleotide sequence of the coding region in Wernicke-Korsakoff and non-Wernicke-Korsakoff individuals
 JOURNAL J. Biol. Chem. 268 (2), 1397-1404 (1993)
 MEDLINE 93123263
 PUBMED 8419340
 REFERENCE 3 (bases 1 to 2104)
 AUTHORS Singleton, C.K.
 TITLE Direct Submission
 JOURNAL Submitted (11-AUG-1992) C.K. Singleton, Vanderbilt University, Box 1820, Station B, Nashville, TN 37235, USA
 FEATURES Location/Qualifiers
 source
 1. .2106
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="pTKa-1, pTKd-1"
 /cell_line="Hela (nucleotides residues 1-200 only)"
 /tissue_type="frontal cortex, liver"
 /clone_lib="frontal cortex and liver cDNA"
 71. .1942
 /EC_number="2.2.1.1"
 /note="homodimer"
 /codon_start=1
 /evidence="experimental"
 /product="transketolase"
 /protein_id="CAA47919.1"
 /db_xref="GI:37267"
 /db_xref="GOA:P29401"
 /translation="MESYHKPDQKQLQALKDTANLRISSIQTAAGSGHPTSCCSA
 AVIMADVLFHTMYKSDQPNRPNDRFVLSKGHAAPILYAVMAEAGFLAEALLNRK
 ISSLDGHPVPKQAFDTATGSLGGLGACGMAYTKYFDKASRYVYCLLDGGLSE
 GSVWEAMAFASYIKLDNLVAIDLINRLGSDPAPLOHMDIYQKCEAFGWHAIIYDG
 HSEVELKAFQAKHQPTAIATKTFKRGITGVDEKESWGKPLPKMAEQIIQBIYS
 IQQSKKLTAPDAPSDVDIANIRMPSLPSYKVGDKIATRAYGQALAKLGHASDR
 IIALDGTNTSTFSEIFKKEHPDRFIECYIAEQNNVSIAGVCATRNTPVFCSTFAAF
 FTRAPDQIRMAAISININLCSHGVSIGEDGSPQMALEDLAFERSVPTSTVFPSD
 GVTEKAVELAAANTKGICFIRTPENAIYNNEDFQVQAKVLSKDDQVTVIGA
 GVTLEHAAALAEKLLKKNIRVLPDFTKPLDKLILDSARATKGRILTVEDHYVIG
 GIGEAIVSSAVGEPGIVTTHLVANVRPSRGPPELLKMFIDRDAIAQAVRGLITKA"
 ORIGIN
 Query Match 1.3%; Score 49; DB 9; Length 2106;
 Best Local Similarity 100.0%; Pred. No. 5.6e-13;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3714 AAATGAAAAA... 3762
 Db 2044 AAATGAAAAA... 2092
 RESULT 74
 BC063356
 LOCUS Silurana tropicalis cDNA clone MGC:75895 IMAGE:5383371, complete cds.
 DEFINITION Silurana tropicalis (Xenopus tropicalis)
 ACCESSION BC063356
 VERSION BC063356.1 GI:38648981
 KEYWORDS MGC.
 SOURCE Silurana tropicalis
 ORGANISM Silurana tropicalis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
 Xenopodinae; Silurana.
 1 (bases 1 to 2177)
 AUTHORS Klein, S.L., Strausberg, R.L., Wagner, L., Pontius, J., Clifton, S.W. and Richardson, P.
 TITLE Genetic and genomic tools for Xenopus research: The NIH Xenopus initiative
 JOURNAL Dev. Dyn. 225 (4), 384-391 (2002)
 PUBMED 12454917
 REFERENCE 2 (bases 1 to 2177)

A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 60 Row: 0 Column: 9
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

FEATURES

Location/Qualifiers

1..2182
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4953665"
/tissue_type="mammary tumor. C3(1)-Tag model. Infiltrating ductal carcinoma. 5 month old virgin mouse."
/clone_lib="NCI CGAP Mam6"
/lab_host="DH10B"
/note="Vector: PCMV-SPORT6"
<1..2182
/gene="LOC217820"
/db_xref="LocusID:217820"
<1..935
/gene="LOC217820"
/codon_start=3
/product="LOC217820 protein"
/protein_id="AAH27154.1"
/db_xref="GI:20071799"
/db_xref="LocusID:217820"
/translation="VFVQKILVGRNSEIEVGEKNAACNLVNGHVDGPIWGLATH
PSRDFLGAEDGTFLVLDIADKMLNKLHGAARTVCYSPGDMVAIGMKGFEY
LLVSLKLTWKKRDRCAIHDIRPSPDRKYLAVGSSNSVDFYDLTGLTLNRTSYK
DIPSVIOMDSADSRHQLQVSSGCKYRHVVEVPGKHLDVHAAIDRITWATWTSLGD
EVMGWSHAEKADVTCVCSHSGISLVTGDFGKVLKYDFPCPEKAKKRFGLHSP
HVTNIRFTSGDRHVVAGDDCSLFWKCVEMPH"
21..776
/gene="LOC217820"
/note="COG2319; Region: FOG: WD40 repeat [General function
prediction only]"
/db_xref="CDD:COG2319"

gene

CDS

misc_feature

ORIGIN

Query Match 1.3%; Score 49; DB 10; Length 2182;
Best Local Similarity 100.0%; Pred. NO. 5.5e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA...AAA 3762

Db 2130 AAATGAAAAA...AAA 2178

RESULT 76

BC049193

LOCUS

DEFINITION Homo sapiens KIAA1277 protein, mRNA (cdna clone IMAGE:4445131), partial cds.

ACCESSION BC049193

VERSION BC049193.1

KEYWORDS GI:29351610

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 2266)

Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,

Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,

Aitschul,S.P., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,

Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,

Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,

Scapleton,M., Soares,M.B., Donald,M.P., Casavant,T.L.,

Scheetz,T.E., Brownstein,M.J., Ussidi,T.B., Toshikuni,S.,

Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,

Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettman,M., Madan,A., Rodrigues,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smallos,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J., and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
1247932
2 (bases 1 to 2266)
Strausberg,R.
Direct Submission
Submitted (21-MAR-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@nsl.nih.gov
Akhtar,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
Blakesley,R.W., Bouffard,G.G., Green,K., Brinkley,C., Brooks,S.,
Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
Hansen,N., Ho,S.-L., Karlens,E., Kwong,P., Legaspi,R.,
Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C.,
McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,
Tsurgenev,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
Young,A., Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 93 Row: c Column: 21.

FEATURES

source

1..2266
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4445131"
/tissue_type="Liver, adenocarcinoma"
/clone_lib="NIH MGC_90"
/lab_host="DH10B"
/note="Vector: PCMV-SPORT6"
<1..2266
/gene="KIAA1277"
/db_xref="LocusID:27147"
<1..1918
/codon_start=2
/product="KIAA1277 protein"
/protein_id="AAH49193.1"
/db_xref="GI:29351611"
/db_xref="LocusID:27147"
/translation="DILDPPMKENYEDIELHRCGLGKVKVNFPPSPSSIPDTLTK
QSLKPAFFQNSERNRNFLLTRKLSRGTGSPKISPPSTSSDDIFFNLGDQON
GRKRIKPLVLRINAIYEVRRGKVRKLSQSMNSGKVTDESDTDEKKA
HSRLNVKSLRQAPRPSLARELIEYQERQLFVYVSVLHKQAGAYVELTQQ
FPLKRSFKLMEAEADQLKAIQFCFPDAKWPVQQTSETSFVLTGEDSRRFQ
YCRLLPGGKRLPEVYCIYSLGCFSLFRILDEVEKRGISPLVQPLMSVMEA
PPPALGKTLVKNFLPGSGTEVIELCRPLDSRLHEVDFESLSVRLVCFVASIL
LERVITADKLKLSKCCAMVALYPPAWQHTYIPVLPANWDIVCSPTPLIGLL

gene

CDS

<p> polyA signal polyA_site ORIGIN </p>	<p> Query Match Best Local Similarity Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0; </p>	<p> 1.3%; Score 49; DB 9; Length 2270; 100.0%; Pred. No. 5.5e-13; 0; Gaps 0; </p>	<p> DH10B, sites NotI + SalI /dev_stage="adult" 2207 </p>
<p> QY Db </p>	<p> 3714 AAATGAAA 3762 2209 AAATGAAA 2257 </p>		
<p> RESULT 78 BC064017 LOCUS DEFINITION ACCESSION VERSION KEYWORDS SOURCE ORGANISM </p>	<p> 2389 bp mRNA linear Mus musculus cdna clone IMAGE:3995710, partial cds. BC064017.1 GI:39794122 Mus musculus (house mouse) Mus musculus Eukaryota; Metazoa; Chordata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 2389) Strausberg R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Hutchins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Spaplen, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Schaeetz, T.E., Brownstein, M.J., Ustin, T.B., Toshikiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abrams, R.D., Mullah, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Munz, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalios, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002) </p>	<p> 2389 bp mRNA linear Mus musculus cdna clone IMAGE:3995710, partial cds. BC064017.1 GI:39794122 Mus musculus (house mouse) Mus musculus Eukaryota; Metazoa; Chordata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 2389) Strausberg R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Hutchins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Spaplen, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Schaeetz, T.E., Brownstein, M.J., Ustin, T.B., Toshikiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abrams, R.D., Mullah, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Munz, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalios, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002) </p>	
<p> JOURNAL PUBMED AUTHORS TITLE JOURNAL </p>	<p> 2 (bases 1 to 2389) Strausberg, R. Direct Submission Submitted (11-DEC-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA NHL-MGC Project URL: http://mgc.nci.nih.gov Contact: MGC help desk Email: gcgbs-x@mail.nih.gov Tissue Procurement: Gilbert Smith, Ph.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) cDNA sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada info@bcgsc.bc.ca Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Neas, Pawan Pandoh, Anna-Liisa Prabhur, Parvaneh Saeedi, Jacqueline Schein, Duane Smalios, Michael Smith, Lorraine Spence, Jeff Scott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra. </p>		

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAK Plate: 131 Row: g Column: 22
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 27370269.

FEATURES

Location/Qualifiers
 1. .2389
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="Czech II"
 /db_xref="taxon:10090"
 /clone="IWAGS:3995710"
 /tissue_type="Mammary tumor metastasized to lung. Tumor arose spontaneously from a senescent normal mammary (clonal) outgrowth infected with the virus MMIV."
 /clone_lib="NCI CGAP Lu29"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"

ORIGIN

Query Match 1.3%; Score 49; DB 10; Length 2389;
 Best Local Similarity 100.0%; Pred. No. 5.5e-13;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA 3762
 |||||
 Db 2339 AATGAAAAA 2387

RESULT 79

BC053747

LOCUS

DEFINITION Mus musculus solute carrier family 43, member 1, mRNA (cdna clone MGC:59491 IMAGE:6330614), complete cds.

ACCESSION

BC053747

VERSION

BC053747.1

KEYWORDS

MGC

SOURCE

Mus musculus (house mouse)

ORGANISM

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 2507)

REFERENCE

AUTHORS

Straussberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Udén, T.B., Toshiyuki, S., Carninci, P., Prange, C., Rana, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shvachenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Smalios, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL

MEDLINE

PUBMED

12477932

REFERENCE

2 (bases 1 to 2507)

Direct Submission

AUTHORS

Straussberg, R.

TITLE

Submitted (13-JUN-2003) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

REMARK

COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Mark Maconochie, Ph.D. and Nancy L. Freeman, Ph.D.
 cDNA Library Preparation: ResGen, Invitrogen Corp
 cDNA library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada
 info@cgsc.bc.ca
 Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Nees, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smalios, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAK Plate: 109 Row: b Column: 18
 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

FEATURES

Location/Qualifiers

1. .2507
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="MGC:59491 IMAGE:6330614"
 /tissue_type="Embryo, day 9 mouse (CS7BL/6 background) otocysts"
 /clone_lib="NIH MGC_130"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6.1"
 1. .2507
 /gene="Slc43a1"
 /note="synonyms: PB39, R00504"
 /db_xref="LocusID:72401"
 /db_xref="MGI:1931352"
 85. .1779
 /codon_start=1
 /product="Slc43a1 protein"
 /protein_id="AAH53747.1"
 /db_xref="GI:31753064"
 /translation="MAPTLQAYRRRMWMACTAVVENLFFSAVLGWSLLIMLKKEG FYSLCPAENRTNTQDQHWTSQDOEKMLNLTGTSFLLSATTLPLGLMDRFG PPRLVGSAFACAACTLMLASRDTEVLPLFLALSLNGFAGTCLTFLITLPNMF GNLRSTFMAIMGVSASSAITFPGLIYDAGVPTVMTFTMWSGLACILFLNCALNWP ABAPAPPEVYTKKILIGALDQKVTGDRFVTHVTVGORLSQKSPSLREGADAFI SSPDIPGTSSETPEKSVFRKSLGSPFLMSLVYMGNTQLRVYMGAMNKILFIVT GKKERTNEQRQKVEETVEFYSSIFGVNQLLCLUTCPILGIMDMWRKDCVDAPTEGT LNNASFGDARDGASTKFTPRYKVKQLTNAINFATLTNILLVGFGLACILKHLQ LIAFLVHTIVRGFFSHSACGLYAAVFPNSHFTGLTGLQSLISAVFALLQQLFKAMVG PLHGDPFWNLGLLLSLFLGLFLLPSLYYYRSRLQREYATNLVDPQKVLNYSKVAT"

gene

CDS

ORIGIN

Query Match 1.3%; Score 49; DB 10; Length 2507;
 Best Local Similarity 100.0%; Pred. No. 5.5e-13;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA 3762

Db 2324 AATGAAAAA 2372

RESULT 80

BC053126

LOCUS

DEFINITION

Danio rerio cDNA clone MGC:63875 IMAGE:6789528, complete cds.

ACCESSION

BC053126

```

VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS

BC053126.1 GI:31419251
MGC.
Danio rerio (zebrafish)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 2582)
Strausberg,R.L., Feingold,B.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altechul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullaly,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whitting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalish,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 2582)
Strausberg,R.
Direct Submission
Submitted (02-JUN-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabs@mail.nih.gov
Tissue Procurement: Leonard I. Zon, M.D.
CDNA Library Preparation: Invitrogen Corp
DNA sequencing by: The I.M.A.G.E. Consortium (LLNL)
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www.shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 117 Row: b Column: 5
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, similarity but not identity to protein.
FEATURES
source
1. .2582
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="MGC:63875 IMAGE:6789528"
/cisue_type="Embryo, zebrafish 24 hpf"
/clone_lib="NCI_CGAP_Zemb2"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6.1"
183. .2096
/codon_start=1
/product="Unknown (protein for MGC:63875)"
/protein_id="AAH53126.1"
/db_xrefs="GI:31419252"
/translation="MNAATAGSYPMASLVGLDHPDITAMLYKFSFAGPVLISRVIC"

CDS

RDMITRSLGVYVNFQPADARALDTNFDVVKGKPIRIMWSORDPSLRKSGVGVN
FIKNDKSIDNKALYDTFSAGFNILSCVKVCDENGSKGYAFVHEFTQDAARAIKKN
GMLNDRKVFVGRFKSRERAEKAKAEFTNVYIKNFDDMDQRLKLPDKYTKT
LSVKVPTDPTGKSRGFGVSEYKEDANKAVEENGTGLNGKTVFVGAQKMERQAE
LKRKVFQKQERISRYQVNLVYIKNLDITIDDEKLKEKFFSGFTTSKAVMLSEERSK
GGFVCFSSPEBATKAVTEMGRIVGSKPLVVALAQRKEERKAHLITNOMRIAGMRA
MPANALINOFOPAGGVFMPVPAQONRTTYVAPNOLTOVRNPNRWQOQGGGQGFQ
MPSLRQPGPRANIRHLAPNASQOPRGMPGTGAQVGGTGOPMGPBPAMGTTTPRAMP
PKYIATIRNTQFQVVFALQQAQPAHVHQVEELTASMLAAAPQKQKQMLGERLE
PLIQAMHPSLAGKITGMLLEIDNSLHMLWLSHESLRKVEEAVLQHQAKQATQ
KVGVTAAATAATS"
429. .1301
/misc_feature
/note="COG0724; Region: RNA-binding proteins (REM domain)
[General function prediction only]"
/db_xref="CDD:COG0724"
1809. .2021
/misc_feature
/note="PABP; Region: Poly-adenylate binding protein,
unique domain"
/db_xref="CDD:pfam00658"

ORIGIN
Query Match 1.3%; Score 49; DB 5; Length 2582;
Best Local Similarity 100.0%; Pred. No. 5.4e-13; Indels 0; Gaps 0;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3714 AAATGAAAAA 3762
Db 2481 AAATGAAAAA 2529

RESULT 81
AX474240 2713 bp DNA linear PAT 12-AUG-2002
LOCUS
DEFINITION Sequence 1 from Patent EP1223218.
AX474240
ACCESSION AX474240
VERSION AX474240.1 GI:22213873
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarhini; Hominiidae; Homo.
REFERENCE
1
AUTHORS Praser,C.C.
TITLE C42000 and cd2001 molecules and uses thereof
JOURNAL Patent: EP 1223218-A 1 17-JUL-2002;
Millennium Pharmaceuticals, Inc. (US)
FEATURES
source
1. .2713
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 1.3%; Score 49; DB 6; Length 2713;
Best Local Similarity 100.0%; Pred. No. 5.4e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3714 AAATGAAAAA 3762
Db 2650 AAATGAAAAA 2698

RESULT 82
AX474270 2713 bp DNA linear PAT 12-AUG-2002
LOCUS
DEFINITION Sequence 31 from Patent EP1223218.
AX474270
ACCESSION AX474270
VERSION AX474270.1 GI:22213882
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarhini; Hominiidae; Homo.

```


[illegible]

QY 3714 AAATGAAAAA..... 3762
 DB 2802 AAATGAAAAA..... 2850

RESULT 90
 AR216298/c
 LOCUS AR216298 2851 bp DNA linear PAT 25-SEP-2002
 DEFINITION Sequence 3 from patent US 6410714.
 ACCESSION AR216298
 VERSION AR216298.1 GI:23314768
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 2851)
 AUTHORS Weber, E.R. and McCall, C.A.
 TITLE Canine low affinity IGE receptor (CD23) nucleic acid molecules and uses thereof
 JOURNAL Patent: US 6410714-A 3 25-JUN-2002;
 FEATURES Location/Qualifiers
 source 1..2851
 /organism="unknown"
 /mol_type="genomic DNA"

ORIGIN
 Query Match 1.3%; Score 49; DB 6; Length 2851;
 Best Local Similarity 100.0%; Pred. No. 5.4e-13;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3714 AAATGAAAAA..... 3762
 DB 50 AAATGAAAAA..... 2

RESULT 91
 BC060912
 LOCUS BC060912 2954 bp mRNA linear VRT 04-NOV-2003
 DEFINITION Danio rerio cDNA clone IMAGE:4144592, partial cds.
 ACCESSION BC060912
 VERSION BC060912.1 GI:38173866
 KEYWORDS
 SOURCE Danio rerio (zebrafish)
 ORGANISM Danio rerio

REFERENCE
 AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullah, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shewchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzyzanski, M.I., Skalska, U., Smalilus, D.E., Scherch, A., Schein, J.E., Jones, S.J. and Matra, M.A.
 Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 22386257
 12477932
 REFERENCE 2 (bases 1 to 2954)
 AUTHORS Strausberg, R.

TITLE Direct Submission
 JOURNAL Submitted (03-NOV-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
 COMMENT Contact: MGC help desk
 Email: cgabbs@mail.nih.gov
 Tissue Procurement: Susan E. Brockerhoff, University of Washington
 cDNA Library Preparation: Susan E. Brockerhoff, University of Washington
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: http://www-shgc.stanford.edu
 Contact: (Dickson, Mark) mcd@paxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.
 Clon distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
 Series: IRAC Plate: 136 Row: b Column: 8.
 FEATURES Location/Qualifiers
 source 1..2954
 /organism="Danio rerio"
 /mol_type="mRNA"
 /strains="wild-type"
 /db_xref="taxon:7955"
 /clone="IMAGE:4144592"
 /tissue_type="Eye, adult retina, 1-2 year old, mixed sex"
 /clone_lib="Zebrafish adult retina cDNA"
 /lab_host="DH10B"
 /notes="Vector: pBluescript SK-"
 <1..1333
 /codon_start=2
 /product="Unknown (protein for IMAGE:4144592)"
 /protein_id="AAH60912.1"
 /db_xref="GI:38173867"
 /translation="SLLPSFCFLDSESSKHTSTWEIDTLGAFDRDFEKKGEKESCPIDQLSHVAKTGTMTQWQKSYFLKLEKVMDDFKASCPQGRGPANPNVYIPPEMKQRLIKVNGYNGIDPTQRLCELLTEPRNVAGTDKFLRGVKNWVYSCVYPTESKGSQCVNRNGVWFPGNTSAPDRNVNGPGTPELRNRPKHSLSNVANGLPDS TESKEQASQESRTVNESASAEASHKRDDEDAATHAEKRLKFDKEE EEDDEDEDEDEIKELDEPHSCTSVASSSDVPSQSTVATVAKDDQDEPSITQ SEVENVGVRSTSESDSPDPKATGESDPKQAESESESAEQTEETNDPVSS SSNNSSDEGVSSAETPSAPSSSTELPAEGSVTBITTSNSETADNMEQD"

ORIGIN
 Query Match 1.3%; Score 49; DB 5; Length 2954;
 Best Local Similarity 100.0%; Pred. No. 5.4e-13;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3714 AAATGAAAAA..... 3762
 DB 2905 AAATGAAAAA..... 2953

RESULT 92
 BC029643
 LOCUS BC029643 3049 bp mRNA linear ROD 11-DEC-2003
 DEFINITION Mus musculus RIKEN cDNA 2810408E11 gene, mRNA (cdna clone MGC:25768 IMAGE:4010810), complete cds.
 ACCESSION BC029643
 VERSION BC029643.1 GI:20987279
 KEYWORDS
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE
 AUTHORS Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus. 1 (bases 1 to 3049)
 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,

Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Scapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fane, J., Heiton, E., Kettelman, M., Madan, A., Rodrigues, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Buckford, G.C., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickinson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Small, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 3049)

Strausberg, R.

Direct Submission

Submitted (06-MAY-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)

DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center

Center code: BCM-HGSC

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>

Contact: amg@bcm.tmc.edu

Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H., Kowis, C.R., Sneed, A.U., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>

Series: IPAK Plate: 30 Row: i Column: 17

This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Similarity but not identity to protein.

FEATURES

source

1. 3049
/organism="Mus musculus"
/mol_type="mRNA"
/strain="Czech II"
/db_xref="taxon:10090"
/clone="MGC:25768 IMAGE:4010810"
/tissue type="Mammary tumor metastasized to lung."
MMTV-LTR/Wnt1 model. Expression driven by an MMTV-LTR enhancer."
/clone lib="NCI CGAP_Lu30"
/lab host="DH10B"
/note="Vector: PCMV-SF0RT6"
1. 3049
/gene="2810408E11Rik"
/db_xref="LocusID:106200"
/db_xref="MGI:1923620"
24 - 2870
/codon_start=1
/product="expressed sequence A147833"
/protein id="AAH29643.1"
/db_xref="GI:20987280"
/db_xref="LocusID:106200"
translation="MSECGRGGGGNNSEDAEBGGPGKSGSLSPAGAAASLEGRI
RGLGASIMARQRELLCGAVALGALLFKFTCSRAKDVIIPAKPPVFFSRRSP
VLDLFGQLDYADHVRQDSEVVVLFYAPWCGQSIARAEIEQAASRLSDQVLVAIN

gene

CDS

CMWNGKCRKQKHFFYPVHLHYRSPGPIEYKGPMSAVYIEKFRVRAKMLLYIPSQ
SALUDFLSSYVERGVLPFGSPQPPGYLTFFTSALHSLKDKYLVGTVRFVGIYDKHL
ARLVSLVSGSVLHREHTSLVFPREVMFTAEINIYKASENOETLFRVLOPHGGKS
LLLNNEKXGPALEFLFPFPLARHLLDEITFEVALEVNCHGDOVERLLOHREV
PSSVSVQCFEFMAALDSFYLKEQFIVHVSGSIECSNPLTSVCSLCVNOQTVGGTR
SRGASFTGSENVLTAPAFEFSLKSCATAPSSPIIEENRYRFPQGLTSTAF
GLSCTRNTLNLYLDNLFVAERLGAAPSAPVKFPAFIVDKESHVILDKQAL
MYTFLESIONFVLYSLPKRHLTGSDSAQFPPTOHLITEVTTDTFWEVTURKQDVLL
YTWQCFPCSLNHFILQARLLPDTFVARIDVSONDLFPWFEMVDRKLPVLPFCN
KDLVSKYVPGDLPTILNLFILHSDAASAPDGPISPTQDCVOSKAVLOREHLS
HVENAMKRSMSLSLARTQVQVRELLSARDGRLLRQRTLEQGHLLRHSOKL
QALYIKKARELOELASGTPPEHTWLKILVATMERELECGGGAKEPAPLKGARPH
SKSNGTTLQPGDTPPTSTSLASTKENTD"

misc_feature

357..644
/note="thiored; Region: Thioredoxin. Thioredoxins are
small enzymes that participate in redox reactions, via the
reversible oxidation of an active centre disulfide bond.
Some members with only the active site are not separated
from the noise"
/db_xref="CDD:pfam00085"
2007..2327

misc_feature

/note="thiored; Region: Thioredoxin. Thioredoxins are
small enzymes that participate in redox reactions, via the
reversible oxidation of an active centre disulfide bond.
Some members with only the active site are not separated
from the noise"
/db_xref="CDD:pfam00085"

ORIGIN

Query Match 1.3%; Score 49; DB 10; Length 3049;
Best Local Similarity 100.0%; Pred. No. 5.4e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762

Db 2971 AAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3019

RESULT 93

AB063100 3072 bp mRNA linear PRI 13-JUN-2001
LOCUS Macaca fascicularis brain cDNA clone:Qt-A-14469, full insert
DEFINITION sequence.

ACCESSION AB063100.1 GI:14388574

VERSION AB063100.1 GI:14388574

KEYWORDS oligo capping, fig (full insert sequence).

SOURCE Macaca fascicularis (crab-eating macaque).

ORGANISM Macaca fascicularis

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;

TITLE Cercopitheciae; Macaca.

JOURNAL 1 (sites)

REFERENCE 2 (bases 1 to 3072)

AUTHORS Osada, N., Hida, M., Kusuda, J., Tanuma, R., Iseki, K., Hirai, M.,

TITL Terao, K., Suzuki, Y., Sugano, S. and Hashimoto, K.

Isolation of full-length cDNA clones from macaque brain CDNA

libraries

Unpublished

JOURNAL 2 (bases 1 to 3072)

AUTHORS Hashimoto, K., Osada, N., Hida, M., Kusuda, J. and Sugano, S.

TITLE Direct Submission

JOURNAL Submitted (11-JUN-2001) Katsuyuki Hashimoto, National Institute of

Infecious Diseases, Division of Genetic Resources, 23-1, Toyama

1-chome, Shinku-ku, Tokyo 162-8640, Japan

(E-mail: khashi@nih.go.jp, URL: <http://www.nih.go.jp/yoken/genebank/>,

Tel: 81-3-5285-1111 (ex.2120), Fax: 81-3-5285-1181)

Lab host: TOP10

Vector: pME18S-PL3 (Acc.No. AB009864)

R. Site1: Drall (CACGTGTGTG)

R. Site2: Drall (CACCATGTG)

Description: 1st strand cDNA was primed with an oligo(dT) primer

[ATGGGCTTTTTTTTTTTT]; double-stranded cDNA was synthesized

using specific 5' and 3' primers and amplified by PCR. The PCR

product was digested with SfiI and size selection was performed to exclude fragments <1.5kb. The SfiI-digested PCR product was cloned into distinct DraIII sites of pME18S-FL3. XhoI sites just outside the DraIII sites can be used to isolate the cDNA insert. Libraries were constructed by oligo-capping method (Sugano et al., , Institute of Medical Science, University of Tokyo).

Custom primer used for sequencing
(5' end primer [CTTCTGCTCTAAAGTGGG] ;
3' end primer [CGACTGCGAGTCGAGCACA]).

FEATURES

source Location/Qualifiers

1..3072

/organism="Macaca fascicularis"

/mol_type="mRNA"

/db_xref="taxon:9541"

/clone="QtrA-14469"

/sex="male"

/tissue_type="temporal lobe right"

/clone_lib="macaque brain cDNA library QtrA"

/dev_stage="adult"

267..1211

CDS

/codon_start=1

/product="hypothetical protein"

/protein_id="BA860805.1"

/db_xref="GI:14388575"

/translation="MGLDSEPGSVLVNVSTALNDTVEFYRTWTSIADKRVNPLMQ
SPWPTLSISTLYLLEFVLPKMKDREPFQMLVLIYFNGVLLNFFIFELPFGSY
NAGYSYIQSDVSNVNEVIRIAALWYFVSKGVEYLDITVFILRKNQVSLHYV
HCTFTLWIGIKVAGQAFPGAGQNSFIHVMSYIGLAAPFMTQKYLWKRVL
TMLQLVQFHTVHTGHTSLYTDCTPPKMWALIAVLSFIFLFIYRTYKPKXP
KTGKTAMGISANGVSKSEKQLVIENGKKQXGKAKGD"

ORIGIN

Query Match 1.3%; Score 49; DB 9; Length 3072;

Best Local Similarity 100.0%; Pred. No. 5.4e-13;

Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA 3762

Db 2991 AAATGAAAAA 3039

RESULT 94

BC054927

LOCUS

BC054927 3124 bp mRNA linear VRT 07-OCT-2003

BC054927 Danio rerio cDNA clone MGC:63655 IMAGE:5612000, complete cds.

ACCESSION

BC054927.1 GI:32766369

VERSION

MGC.

KEYWORDS

Source

ORGANISM

Danio rerio

Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Ostariophysi;

Cypriniformes; Cyprinidae; Danio.

1 (bases 1 to 3124)

Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,

Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,

Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,

Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,

Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,

Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,

Schetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,

Carinci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,

Abramson,R.D., Mullaly,S.J., Bosak,S.A., McEwan,P.J.,

McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,

Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,

Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,

Fahy,J., Helton,E., Kettner,M., Madan,A., Rodriguez,S.,

Sanchez,A., Whitting,M., Madan,A., Young,A.C., Shevchenko,Y.,

Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,

Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,

Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalhus,D.E.,

Schmerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.

TITLE

JOURNAL

MEDLINE

REFERENCE

PUBMED

12477932

2 (bases 1 to 3124)

Strausberg,R.

Direct Submission

TITLE

JOURNAL

Submitted (08-JUL-2003) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

NIH-MGC Project URL: http://mgc.nci.nih.gov

Contact: MGC help desk

Email: cgabbs@mail.nih.gov

Tissue Procurement: Dr. Sumio Sugano

CDNA Library Preparation: Dr. Sumio Sugano

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Genome Sequence Centre,

BC Cancer Agency, Vancouver, BC, Canada

info@bcsc.bc.ca

Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,

Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,

Letitia Hsiao, Martin Krzywinski, Reta Kutache, Oliver Lee, Soo

Sen Lee, Victor Ling, Carrie Mathewson, Candice McEay, Steven

Ness, Pawan Pandon, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline

Schein, Duane Smalhus, Michael Smith, Lorraine Spence, Jeff Stott,

Michael Thorne, Minada Tsai, Natasja van den Bosch, Jill Vardy,

George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov

Series: IRAC Plate: 118 Row: 0 Column: 1

This clone was selected for full length sequencing because it

passed the following selection criteria: Hexamer frequency ORP

analysis, Similarity but not identity to protein.

Location/Qualifiers

source

1..3124

/organism="Danio rerio"

/mol_type="mRNA"

/db_xref="taxon:7955"

/clone="MGC:63655 IMAGE:5612000"

/tissue_type="Whole body, adult male"

/clone_lib="Sugano SJD adult male"

/lab_host="DH10B"

/note="Vector: pME18S-FL3"

222..890

/codon_start=1

/product="Unknown (protein for MGC:63655)"

/protein_id="AAH54927.1"

/db_xref="GI:32766370"

/translation="MDRISRYLSRVYGMHPKDTARQLVLAQDGLVETLVGCKGSK

AGTEQGYWLPQDEMGSEKSPSPSEALPDDWEASHDWYCFCHLPQGVMECD

GCFRVYHLRCVSEDRPRTTSHWQCGICRGRKNLNKQENTYTKFILGRMCRVAV

DLKSFEEFRADQLVHVTAILNGVNSQTEIARLLYNDTCLLASCGFVFCARSPSD

WGPA"

misc_feature

483..617

/notes="PHD; Region: PHD zinc finger"

/db_xref="CDD:smart00249"

627..818

/notes="BROMO; Region: bromo domain"

/db_xref="CDD:smart00297"

ORIGIN

Query Match 1.3%; Score 49; DB 5; Length 3124;

Best Local Similarity 100.0%; Pred. No. 5.3e-13;

Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA 3762

Db 3023 AAATGAAAAA 3071


```

RESULT 95
BC036825      3145 bp      mRNA      linear      PRI 26-AUG-2002
LOCUS        Homo sapiens, clone IMAGE:5191998, mRNA, partial cds.
DEFINITION
ACCESSION    BC036825
VERSION      BC036825.1  GI:22477362
KEYWORDS
SOURCE       Homo sapiens (human)
ORGANISM     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (bases 1 to 3145)
AUTHORS      Strausberg, R.
TITLE        Direct Submission
JOURNAL      Submitted (23-AUG-2002) National Institutes of Health, Mammalian
              Gene Collection (MGC), Cancer Genomics Office, National Cancer
              Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
              USA
REMARK       NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT      Contact: MGC help desk
              Email: cgapbs-remail.nih.gov
              Tissue Procurement: Life Technologies, Inc.
              cDNA Library Preparation: Life Technologies, Inc.
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
              DNA Sequencing by: National Institutes of Health Intramural
              Sequencing Center (NISC),
              Gaithersburg, Maryland;
              Web site: http://www.nisc.nih.gov/
              Contact: nisc.mgc@ngri.nih.gov
              Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
              Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
              Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
              Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Lalic, P., Legaspi, R.,
              Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
              McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
              Tsurgou, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
              Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/ILNL at: http://image.llnl.gov
Series: IRAP Plate: 78 Row: e Column: 23
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis.

FEATURES             Location/Qualifiers
     source           1..3145
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5191998"
                     /tissue_type="Brain, adult, 6 pooled whole brains"
                     /clone_lib="NIH MGC_114"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     CDS              1..692
                     /codon_start=3
                     /product="Unknown (protein for IMAGE:5191998)"
                     /protein_id="AAH36825.1"
                     /db_xref="GI:22477363"
                     /translation="CLOHSPQSQLCTLLCWLGRPPFRKQSTSPGLEGVPRF
                     LKRLLLVGNLSSVWAVQVDIFTAGKLALALIIIMGIQVICKGYFVLEPKNAFE
                     NFQEPDGLWALFQGSFAYGWNFLNVTEELVDVFKLFFAGAREGHLPSPVLAMTH
                     VKRCTPIALFTCTICITLLMLVTSMDYTLINVGFINVLYFGVTVAGQIVLRWKKPDI
                     FPIKPLPIQG"

ORIGIN
Query Match      1.3%; Score 49; DB 9; Length 3145;
Best Local Similarity 100.0%; Pred. No. 5.3e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3714 AAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
        |||||||
        DB      3060 AAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3108

FEATURES             Location/Qualifiers
     source           1..3145
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5191998"
                     /tissue_type="Brain, adult, 6 pooled whole brains"
                     /clone_lib="NIH MGC_114"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     CDS              1..692
                     /codon_start=3
                     /product="Unknown (protein for IMAGE:5191998)"
                     /protein_id="AAH36825.1"
                     /db_xref="GI:22477363"
                     /translation="CLOHSPQSQLCTLLCWLGRPPFRKQSTSPGLEGVPRF
                     LKRLLLVGNLSSVWAVQVDIFTAGKLALALIIIMGIQVICKGYFVLEPKNAFE
                     NFQEPDGLWALFQGSFAYGWNFLNVTEELVDVFKLFFAGAREGHLPSPVLAMTH
                     VKRCTPIALFTCTICITLLMLVTSMDYTLINVGFINVLYFGVTVAGQIVLRWKKPDI
                     FPIKPLPIQG"

ORIGIN
Query Match      1.3%; Score 49; DB 9; Length 3145;
Best Local Similarity 100.0%; Pred. No. 5.3e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3714 AAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
        |||||||
        DB      3060 AAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3108

```

```

RESULT 96
HSM800211      3174 bp      mRNA      linear      PRI 18-FEB-2000
LOCUS        Homo sapiens mRNA; cDNA DKFZp586C241 (from clone DKFZp586C241).
DEFINITION
ACCESSION    AL049426
VERSION      AL049426.1  GI:4500206
KEYWORDS
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (bases 1 to 3174)
AUTHORS      Wambutt, R., Heubner, D., Mewes, H.W., Gassenhuber, J. and Wiemann, S.
TITLE        Direct Submission
JOURNAL      Submitted (10-MAR-1999) MIPS, Am Klopferspitz 18a, D-82152
              Martinsried, GERMANY
COMMENT      Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
              Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
              sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing
              consortium of the German Genome Project.
              This clone (DKFZp586C241) is available at the RZPD in Berlin.
              Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
              Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
              information about the clone and the sequencing project is available
              at http://www.mips.biochem.mpg.de/proj/cDNA/.

FEATURES             Location/Qualifiers
     source           1..3174
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="RZPD:DKFZp586C241"
                     /db_xref="taxon:9606"
                     /clone="DKFZp586C241"
                     /tissue_type="uterus"
                     /clone_lib="585 (synonym: hute1). Vector pSport1, host
                     DH10B; sites NotI + SalI/MluI"
                     /dev_stage="adult"
     polyA_signal      3104..3109
     polyA_site        3118
     ORIGIN            3118

Query Match      1.3%; Score 49; DB 9; Length 3174;
Best Local Similarity 100.0%; Pred. No. 5.3e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3714 AAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
        |||||||
        DB      3119 AAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3167

RESULT 97
BC032176      3290 bp      mRNA      linear      ROD 16-APR-2003
LOCUS        Mus musculus RIKEN cDNA 953006B08 gene, mRNA (cDNA clone
DEFINITION    IMAGE:5322259), partial cds.
ACCESSION    BC032176
VERSION      BC032176.1  GI:21595134
KEYWORDS
SOURCE       Mus musculus (house mouse)
ORGANISM     Mus musculus
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE    1 (bases 1 to 3290)
AUTHORS      Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
              Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
              Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
              Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
              Diatchenko, L., Marusik, K., Farmer, A.A., Rubin, G.M., Hong, L.,
              Sapleten, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
              Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
              Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J.,
              Abramson, R.D., Mullen, S.J., Bosak, S.A., McEwan, P.J.,

```

McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Paray, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, J.S., Krzywinski, M.I., Skalski, U., Smalinski, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

22388257

12477932

2 (bases 1 to 3290)

Strausberg, R.

Direct Submission

Submitted (06-JUN-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgabs-r@mail.nih.gov

Tissue Procurement: Jeffrey Green M.D.

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland;

Web site: <http://www.nisc.nih.gov/>

Contact: nisc_mgc@hgr.nih.gov

Akter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granate, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, P., Legaspi, R., Maduro, Q.L., Mastello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W., Tsugeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAC Plate: 56 Row: 1 Column: 17

This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

FEATURES

source

1. .3290

/organism="Mus musculus"

/mol_type="mRNA"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone="IMAGE:5322259"

/tissue_type="Mammary tumor, C3(1)-Tag model. Infiltrating ductal carcinoma. 5 month old virgin mouse."

/clone_lib="NCI_CGAP_Mame"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

<1. .3290

/gene="9530006B08Rik"

/db_xref="LOCUSID:229584"

/db_xref="MGI:2442117"

<1. .2696

/gene="9530006B08Rik"

/codon_start=3

/product="9530006B08Rik protein"

/protein_id="AAH32176.1"

/db_xref="GI:21595135"

/db_xref="LOCUSID:229584"

/db_xref="MGI:2442117"

translation="TRPLDQNGVDGHTICOHYQRPSTFFQLOCHLNVHVSPEST TKCKICENAFSEPLFLOKMDTHKPGMPYVCQVCRSLYSYSEVDVHFRMIHEDTR

ORIGIN

Query Match 1.3%; Score 49; DB 10; Length 3290;

Best Local Similarity 100.0%; Pred. No. 5.3e-13;

Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAA 3762

DB 3242 AAATGAAA 3290

RESULT 98

BC034128

LOCUS

DEFINITION

Mus musculus, Similar to neuronal specific transcription factor DAT1, clone IMAGE:4913908, mRNA, partial cds.

ACCESSION

BC034128

VERSION

BC034128.1 GI:21706632

KEYWORDS

Mus musculus (house mouse)

SOURCE

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus; Strausberg, R.

REFERENCE

1 (bases 1 to 3387)

AUTHORS

Strausberg, R.

TITLE

Direct Submission

JOURNAL

Submitted (02-JUL-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgabs-r@mail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.

CDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)

Center, Stanford University School of Medicine, Stanford, CA 94305

Web site: <http://www.shgc.stanford.edu>

Contact: (Dickson, Mark) mcdpaxil.stanford.edu

Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAC Plate: 44 Row: 1 Column: 3

This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

FEATURES

source

1. .3387

/organism="Mus musculus"

/mol_type="mRNA"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone="IMAGE:4913908"

/tissue_type="Salivary gland, 10 week old female mouse"

/clone_lib="NCI_CGAP_SG2"

/lab_host="DH10B"

Search completed: April 23, 2004, 08:22:46
Job time : 14172 secs

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE AUTHORS

1 (bases 1 to 3608)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenman, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.N., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smailus, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
12477932
2 (bases 1 to 3608)
Strausberg, R.
Direct Submission
Submitted (01-MAY-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunaratne, P.R., Garcia, A.M., Lu, X., Hulyk, S.W., Loulsegad, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.N., Nanavati,
A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAC Plate: 37 Row: d Column: 20
This clone was selected for full length sequencing because it
passed the following selection criteria: Similarity but not
identity to protein.

FEATURES

Location/Qualifiers
1..3608
/organism="Mus musculus"
/mol_type="mrna"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4163231"
/tissue_type="Liver, normal. 5 month old male mouse."
/clone_lib="NCI CGAP_Li9"
/lab_host="DH108"
/note="Vector: PCMV-SPORT6"

ORIGIN

Query Match 1.3% Score 49; DB 10; Length 3608;
Best Local Similarity 100.0%; Pred. No. 5.3e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3714 AAATGAAA 3762
|||||
Db 3548 AAATGAAA 3596

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 23, 2004, 03:54:40 ; Search time 244 Seconds
(without alignments) 8556.251 Million cell us

Title: US-10-005-907-1
 Perfect score: 3762
 Sequence: 1 qagaaacggagtactctga:.....aaaaaaaaaaaaaaaaaaaa 3762

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 682709 seqs, 277475446 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

MAXIMUM DB REF TEMPS: 2000000000

[illegible]

```
Database :
1: /cgm2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgm2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgm2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgm2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgm2_6/ptodata/2/ina/PTCUS_COMB.seq.*
6: /cgm2_6/ptodata/2/ina/backfiles1.seq.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query			ID	Description
	Score	Match	Length		
1	49	1.3	159	US-09-621-976-17182	Sequence 17182, A
2	49	1.3	882	US-08-909-965C-9	Sequence 9, Appli
3	49	1.3	1933	US-09-920-759-3	Sequence 3, Appli
4	49	1.3	2851	US-09-535-521-1	Sequence 1, Appli
5	49	1.3	2851	US-09-535-521-3	Sequence 3, Appli
6	48	1.3	191	US-09-621-976-9448	Sequence 9448, Ap
7	48	1.3	619	US-09-489-847-558	Sequence 58, Appli
8	48	1.3	1544	US-09-187-993-14	Sequence 14, Appli
9	47	1.2	105	US-08-735-381-2	Sequence 2, Appli
10	47	1.2	105	US-09-183-619-1	Sequence 1, Appli
11	47	1.2	105	US-09-201-674-2	Sequence 2, Appli
12	47	1.2	165	US-09-621-976-13195	Sequence 13195, A
13	47	1.2	266	US-09-621-976-16813	Sequence 16813, A
14	47	1.2	326	US-09-621-976-18024	Sequence 18024, A
15	47	1.2	329	US-09-621-976-18012	Sequence 18012, A
16	47	1.2	332	US-09-621-976-16031	Sequence 16031, A
17	47	1.2	332	US-09-621-976-16050	Sequence 16050, A
18	47	1.2	332	US-09-621-976-16053	Sequence 16053, A
19	47	1.2	333	US-09-621-976-16032	Sequence 16032, A
20	47	1.2	333	US-09-621-976-16045	Sequence 16045, A
21	47	1.2	334	US-09-621-976-16044	Sequence 16044, A
22	47	1.2	335	US-09-621-976-16061	Sequence 16061, A
23	47	1.2	336	US-09-621-976-16013	Sequence 16013, A
24	47	1.2	338	US-09-621-976-16041	Sequence 16041, A
25	47	1.2	347	US-09-621-976-16026	Sequence 16026, A
26	47	1.2	357	US-09-621-976-16058	Sequence 16058, A
27	47	1.2	358	US-09-621-976-16008	Sequence 16008, A

101 46 1.2 1630 4 US-10-162-012-39
102 46 1.2 1699 4 US-09-152-060-19
103 46 1.2 1713 4 US-09-000-062-1
104 46 1.2 1713 4 US-08-945-144A-1
105 46 1.2 1772 2 US-08-960-022-13
106 46 1.2 1790 4 US-09-866-028-31
107 46 1.2 1804 2 US-08-504-459-5
108 46 1.2 1910 3 US-08-974-631-7
109 46 1.2 1958 4 US-08-663-034-3
110 46 1.2 2013 4 US-09-596-196-3
111 46 1.2 2096 3 US-09-008-481A-10
112 46 1.2 2096 3 US-09-195-666A-16
113 46 1.2 2096 3 US-09-309-592-10
114 46 1.2 2096 4 US-09-635-705-16
115 46 1.2 2096 4 US-09-634-858A-16
116 46 1.2 2096 4 US-08-863-927C-16
117 46 1.2 2205 3 US-08-888-077A-41
118 46 1.2 2323 4 US-09-149-476-24
119 46 1.2 2485 4 US-09-889-463A-9
120 46 1.2 2674 4 US-09-817-180-1
121 46 1.2 2674 4 US-10-003-295-1
122 46 1.2 2744 3 US-09-071-101-1
123 46 1.2 2744 3 US-09-369-618-1
124 46 1.2 2744 3 US-09-369-617-1
125 46 1.2 2822 4 US-09-907-794A-48
126 46 1.2 2822 4 US-09-905-125A-48
127 46 1.2 2822 4 US-09-903-775A-48
128 46 1.2 3138 1 US-07-867-106-4
129 46 1.2 3441 2 US-08-742-783-1
130 46 1.2 3850 4 US-10-072-094-94
131 46 1.2 3722 4 US-10-164-595-9
132 46 1.2 3862 4 US-10-164-595-5
133 46 1.2 3871 2 US-08-599-455B-3
134 46 1.2 3871 3 US-09-069-781B-3
135 46 1.2 3871 4 US-09-137-132-3
136 46 1.2 3871 4 US-08-864-564A-3
137 46 1.2 3871 4 US-09-094-410-3
138 46 1.2 3871 4 US-08-708-123D-3
139 46 1.2 3871 4 US-08-583-153A-3
140 46 1.2 3871 4 US-08-570-142D-3
141 46 1.2 3871 4 US-08-638-524B-3
142 46 1.2 3894 4 US-09-511-625B-3
143 46 1.2 3937 4 US-10-164-595-7
144 46 1.2 3985 4 US-10-164-595-3
145 46 1.2 4456 4 US-09-095-443-1
146 46 1.2 4895 3 US-09-053-866-1
147 46 1.2 4895 4 US-09-479-130-1
148 46 1.2 4895 4 US-09-472-130A-1
149 46 1.2 6671 1 US-08-280-443-1
150 46 1.2 6671 1 US-08-457-459-1

ALIGNMENTS

RESULT 1
US-09-621-976-17182
; Sequence 17182, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET 054PR2
; CURRENT APPLICATION NUMBER: US/09/621, 976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 17182
; LENGTH: 159
; TYPE: DNA
; ORGANISM: Homo sapiens

US-09-621-976-17182

Query Match 1.3%; Score 49; DB 4; Length 159;
Best Local Similarity 100.0%; Pred. No. 3.2e-10;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAA 3762
Db 100 AAATGAAA 148

RESULT 2

US-08-909-965C-9
; Sequence 9, Application US/08909965C
; Patent No. 5936078
; GENERAL INFORMATION:
; APPLICANT: Kuga Tetsuo
; APPLICANT: Nakagawa Satoshi
; APPLICANT: Sakaki Yoshiyuki
; APPLICANT: Zhao Nanding
; APPLICANT: Hashida Hideji
; TITLE OF INVENTION: NOVEL DNA, NOVEL POLYPEPTIDE
; TITLE OF INVENTION: AND NOVEL ANTIBODY
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FITZPATRICK, CELLA, HARPER AND SCINTO
; STREET: 277 Park Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10172-0194
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/909,965C
; FILING DATE: August 12, 1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 322745/95
; APPLICATION NUMBER: PCT/JP96/03630
; FILING DATE: 12-Dec-1996
; FILING DATE: 12-Dec-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Lawrence S. Perry
; REGISTRATION NUMBER: 31865
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-2400
; TELEFAX: 212-758-2982
; TELEX: 236282
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 882 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; ORIGINAL SOURCE:
; ORGANISM: human
; IMMEDIATE SOURCE:
; CLONE: F1180
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 97 to 195
; IDENTIFICATION METHOD: by experiment
; US-08-909-965C-9

Query Match 1.3%; Score 49; DB 2; Length 882;
Best Local Similarity 100.0%; Pred. No. 2.8e-10;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AATGAA 3762
|||||
Db 825 AATGAA 873

RESULT 3
US-09-920-759-3
; Sequence 3, Application US/09920759
; Patent No. 6537811
; GENERAL INFORMATION:
; APPLICANT: Brenda F. Baker
; APPLICANT: Susan M. Freier
; TITLE OF INVENTION: ANTISENSE MODULATION OF SAP-1 EXPRESSION
; FILE REFERENCE: RTS-0267
; CURRENT APPLICATION NUMBER: US/09/920,759
; CURRENT FILING DATE: 2001-08-01
; NUMBER OF SEQ ID NOS: 91
; SEQ ID NO 3
; LENGTH: 1933
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (150)....(1445)
US-09-920-759-3

Query Match 1.3%; Score 49; DB 4; Length 1933;
Best Local Similarity 100.0%; Pred. No. 2.6e-10;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AATGAA 3762
|||||
Db 1878 AATGAA 1926

RESULT 4
US-09-535-521-1
; Sequence 1, Application US/09535521
; Patent No. 6410714
; GENERAL INFORMATION:
; APPLICANT: Weber, Eric R.
; APPLICANT: McCall, Catherine A.
; TITLE OF INVENTION: NOVEL CANINE LOW AFFINITY IGE RECEPTOR (CANINE CD23)
; TITLE OF INVENTION: NOVEL CANINE LOW AFFINITY IGE RECEPTOR (CANINE CD23)
; FILE REFERENCE: AL-5
; CURRENT APPLICATION NUMBER: US/09/535,521
; CURRENT FILING DATE: 2000-03-24
; EARLIER APPLICATION NUMBER: 60/125,913
; EARLIER FILING DATE: 1999-03-24
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2851
; TYPE: DNA
; ORGANISM: Canis familiaris
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (199)..(1077)
US-09-535-521-1

Query Match 1.3%; Score 49; DB 4; Length 2851;
Best Local Similarity 100.0%; Pred. No. 2.5e-10;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AATGAA 3762
|||||
Db 2802 AATGAA 2850

RESULT 5
US-09-535-521-3/c
; Sequence 3, Application US/09535521
; Patent No. 6410714

GENERAL INFORMATION:
; APPLICANT: Weber, Eric R.
; APPLICANT: McCall, Catherine A.
; TITLE OF INVENTION: NOVEL CANINE LOW AFFINITY IGE RECEPTOR (CANINE CD23)
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES AND USES THEREOF
; FILE REFERENCE: AL-5
; CURRENT APPLICATION NUMBER: US/09/535,521
; CURRENT FILING DATE: 2000-03-24
; EARLIER APPLICATION NUMBER: 60/125,913
; EARLIER FILING DATE: 1999-03-24
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 3
; LENGTH: 2851
; TYPE: DNA
; ORGANISM: Canis familiaris
US-09-535-521-3

Query Match 1.3%; Score 49; DB 4; Length 2851;
Best Local Similarity 100.0%; Pred. No. 2.5e-10;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AATGAA 3762
|||||
Db 50 AATGAA 2

RESULT 6
US-09-621-976-9448
; Sequence 9448, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 9448
; LENGTH: 191
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-9448

Query Match 1.3%; Score 48; DB 4; Length 191;
Best Local Similarity 100.0%; Pred. No. 7.9e-10;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAA 3762
|||||
Db 143 AATGAA 190

RESULT 7
US-09-489-847-58
; Sequence 58, Application US/09489847
; Patent No. 6476195
; GENERAL INFORMATION:
; APPLICANT: Rosen et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: PZ031P1
; CURRENT APPLICATION NUMBER: US/09/489,847
; CURRENT FILING DATE: 2000-01-24
; EARLIER APPLICATION NUMBER: PCT/US99/17130
; EARLIER FILING DATE: 1999-07-29
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; EARLIER APPLICATION NUMBER: 60/095,486
; EARLIER FILING DATE: 1998-08-05
; EARLIER APPLICATION NUMBER: 60/096,319

```

CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304-1126
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/735,381
FILING DATE: 21-OCT-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Choi, Wendy A.
REGISTRATION NUMBER: 36,697
REFERENCE/DOCKET NUMBER: 10950427-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-857-4125
TELEFAX: 650-852-8063
TELEX: 348-461
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 105 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-735-381-2

Query Match 1.2%; Score 47; DB 2; Length 105;
Best Local Similarity 100.0%; Pred. No. 2.1e-09;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3716 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
DB 53 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 99

RESULT 10
US-09-183-619-1
; Sequence 1, Application US/09183619
; Patent No. 6103474
; GENERAL INFORMATION:
; APPLICANT: DELLINGER, DOUGLAS J.
; APPLICANT: DAHM, SUEANN C.
; APPLICANT: ILSLEY, DIANE D.
; APPLICANT: ACH, ROBERT A.
; APPLICANT: TROLL, MARK A.
; TITLE OF INVENTION: HYBRIDIZATION ASSAY SIGNAL ENHANCEMENT
; FILE REFERENCE: 10981619-1
; CURRENT APPLICATION NUMBER: US/09/183,619
; CURRENT FILING DATE: 1999-10-30
; EARLIER APPLICATION NUMBER: 08/735,381
; EARLIER FILING DATE: 1996-10-21
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 105
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Target analyte
US-09-183-619-1

Query Match 1.2%; Score 47; DB 3; Length 105;
Best Local Similarity 100.0%; Pred. No. 2.1e-09;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3716 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762

```

Db 53 ATGAA 99

RESULT 11

US-09-201-674-2

Sequence 2, Application US/09201674

Patent No. 611682

GENERAL INFORMATION:

APPLICANT: Dellinger, Douglas J.

Dahn, SueAnn

Troll, Mark

TITLE OF INVENTION: SIGNAL ENHANCEMENT METHOD AND KIT

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:

ADDRESSEE: Hewlett-Packard Company, Legal Dept.,

Intellectual Property

STREET: 1501 Page Mill Road, MS 4U-10

CITY: Palo Alto

STATE: California

COUNTRY: USA

ZIP: 94304-1126

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/201,674

FILING DATE: 30-Oct-1998

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/735,381

FILING DATE: 21-Oct-1996

ATTORNEY/AGENT INFORMATION:

NAME: Choi, Wendy A.

REGISTRATION NUMBER: 36,697

REFERENCE/DOCKET NUMBER: 10950427-1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650-857-4125

TELEFAX: 650-852-8063

TELEX: 348-463

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 105 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid

HYPOTHETICAL: NO

ANTI-SENSE: NO

SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-201-674-2

Query Match 1.2%; Score 47; DB 3; Length 105;

Best Local Similarity 100.0%; Pred. No. 2.1e-09;

Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3716 ATGAAA 3762

Db 53 ATGAAA 99

RESULT 12

US-09-621-976-19195

Sequence 19195, Application US/09621976

Patent No. 6639063

GENERAL INFORMATION:

APPLICANT: Dumas Milne Edwards, J.B.

APPLICANT: Jobert, S.

APPLICANT: Giordano, J.Y.

TITLE OF INVENTION: ESTs and Encoded Human Proteins.

FILE REFERENCE: GENSET.054PR2

CURRENT APPLICATION NUMBER: US/09/621,976

CURRENT FILING DATE: 2000-07-21

NUMBER OF SEQ ID NOS: 19335

SOFTWARE: Patent.pm

SEQ ID NO 16024

LENGTH: 326

TYPE: DNA

ORGANISM: Homo sapiens

US-09-621-976-16024

Query Match 1.2%; Score 47; DB 4; Length 326;

Best Local Similarity 100.0%; Pred. No. 1.9e-09;

Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3716 ATGAAA 3762

Db 202 ATGAAA 248

RESULT 13

US-09-621-976-16813

Sequence 16813, Application US/09621976

Patent No. 6639063

GENERAL INFORMATION:

APPLICANT: Dumas Milne Edwards, J.B.

APPLICANT: Jobert, S.

APPLICANT: Giordano, J.Y.

TITLE OF INVENTION: ESTs and Encoded Human Proteins.

FILE REFERENCE: GENSET.054PR2

CURRENT APPLICATION NUMBER: US/09/621,976

CURRENT FILING DATE: 2000-07-21

NUMBER OF SEQ ID NOS: 19335

SOFTWARE: Patent.pm

SEQ ID NO 16813

LENGTH: 266

TYPE: DNA

ORGANISM: Homo sapiens

US-09-621-976-16813

Query Match 1.2%; Score 47; DB 4; Length 266;

Best Local Similarity 100.0%; Pred. No. 1.9e-09;

Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3716 ATGAAA 3762

Db 202 ATGAAA 248

RESULT 14

US-09-621-976-16024

Sequence 16024, Application US/09621976

Patent No. 6639063

GENERAL INFORMATION:

APPLICANT: Dumas Milne Edwards, J.B.

APPLICANT: Jobert, S.

APPLICANT: Giordano, J.Y.

TITLE OF INVENTION: ESTs and Encoded Human Proteins.

FILE REFERENCE: GENSET.054PR2

CURRENT APPLICATION NUMBER: US/09/621,976

CURRENT FILING DATE: 2000-07-21

NUMBER OF SEQ ID NOS: 19335

SOFTWARE: Patent.pm

SEQ ID NO 16024

LENGTH: 326

TYPE: DNA

ORGANISM: Homo sapiens

US-09-621-976-16024

Query Match 1.2%; Score 47; DB 4; Length 326;

Best Local Similarity 100.0%; Pred. No. 1.9e-09;

Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3716 ATGAAA 3762

Db 202 ATGAAA 248

```
Db 268 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 314

RESULT 15
US-09-621-976-16012
; Sequence 16012, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 16012
; LENGTH: 329
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-16012

Query Match 1.2%; Score 47; DB 4; Length 329;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3716 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 268 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 314

RESULT 16
US-09-621-976-16031
; Sequence 16031, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 16031
; LENGTH: 332
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-16031

Query Match 1.2%; Score 47; DB 4; Length 332;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3716 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 271 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 317

RESULT 17
US-09-621-976-16050
; Sequence 16050, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
```

```
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 16050
; LENGTH: 332
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-16050

Query Match 1.2%; Score 47; DB 4; Length 332;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3716 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 268 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 314

RESULT 18
US-09-621-976-16053
; Sequence 16053, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 16053
; LENGTH: 332
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-16053

Query Match 1.2%; Score 47; DB 4; Length 332;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3716 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 268 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 314

RESULT 19
US-09-621-976-16032
; Sequence 16032, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 16032
; LENGTH: 333
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-16032

Query Match 1.2%; Score 47; DB 4; Length 333;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3716 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 268 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 314
```

```
RESULT 20
US-09-621-976-16045
; Sequence 16045, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 16045
; LENGTH: 333
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-16045
Query Match 1.2%; Score 47; DB 4; Length 333;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3716 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
DB 268 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 314

RESULT 21
US-09-621-976-16044
; Sequence 16044, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 16044
; LENGTH: 334
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; LOCATION: 334
; OTHER INFORMATION: n=a, g, c or t
US-09-621-976-16044
Query Match 1.2%; Score 47; DB 4; Length 334;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3716 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
DB 268 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 314

RESULT 22
US-09-621-976-16061
; Sequence 16061, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 16061
; LENGTH: 338
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
US-09-621-976-16061
Query Match 1.2%; Score 47; DB 4; Length 335;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3716 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
DB 268 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 314

RESULT 23
US-09-621-976-16013
; Sequence 16013, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 16013
; LENGTH: 336
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-16013
Query Match 1.2%; Score 47; DB 4; Length 336;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3716 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
DB 268 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 314

RESULT 24
US-09-621-976-16041
; Sequence 16041, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 16041
; LENGTH: 338
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
US-09-621-976-16041
Query Match 1.2%; Score 47; DB 4; Length 336;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3716 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
DB 268 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 314
```

; LOCATION: 135,197
; OTHER INFORMATION: n=a, g, c or t
US-09-621-976-16041

Query Match 1.2%; Score 47; DB 4; Length 338;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3716 ATGAA 3762
Db 270 ATGAA 316

RESULT 25

US-09-621-976-16026
; Sequence 16026, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 16026
; LENGTH: 347
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 347
; OTHER INFORMATION: n=a, g, c or t
US-09-621-976-16026

Query Match 1.2%; Score 47; DB 4; Length 347;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3716 ATGAA 3762
Db 268 ATGAA 314

RESULT 26

US-09-621-976-16058
; Sequence 16058, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 16058
; LENGTH: 357
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-16058

Query Match 1.2%; Score 47; DB 4; Length 357;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3716 ATGAA 3762
Db 268 ATGAA 314

RESULT 27

US-09-621-976-16008
; Sequence 16008, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 16008
; LENGTH: 359
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-16008

Query Match 1.2%; Score 47; DB 4; Length 359;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3716 ATGAA 3762
Db 268 ATGAA 314

RESULT 28

US-09-621-976-16019
; Sequence 16019, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 16019
; LENGTH: 359
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-16019

Query Match 1.2%; Score 47; DB 4; Length 359;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3716 ATGAA 3762
Db 268 ATGAA 314

RESULT 29

US-09-621-976-16010
; Sequence 16010, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335


```
; SOFTWARE: Patent.pm
; SEQ ID NO 16010
; LENGTH: 362
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-16010

Query Match      1.2%; Score 47; DB 4; Length 362;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3716 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 268 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 314

RESULT 30
US-09-621-976-16042
; Sequence 16042, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621.976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 16042
; LENGTH: 365
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-16042

Query Match      1.2%; Score 47; DB 4; Length 365;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3716 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 268 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 314

RESULT 31
US-09-621-976-16048
; Sequence 16048, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621.976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 16048
; LENGTH: 371
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-16048

Query Match      1.2%; Score 47; DB 4; Length 371;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3716 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 269 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 315

; SOFTWARE: Patent.pm
; SEQ ID NO 16010
; LENGTH: 362
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-16010

Query Match      1.2%; Score 47; DB 4; Length 362;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3716 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 268 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 314

RESULT 32
US-09-105-542A-2
; Sequence 2, Application US/09105542A
; Patent No. 6323329
; GENERAL INFORMATION:
; APPLICANT: Bullerdiel, Jorn
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES OF GENES OF THE
; FILE REFERENCE: BOHWA.001C1CP
; CURRENT APPLICATION NUMBER: US/09/105.542A
; CURRENT FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: PCT/DE96/02494
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 615
; TYPE: DNA
; ORGANISM: Homo Sapien
US-09-105-542A-2

Query Match      1.2%; Score 47; DB 4; Length 615;
Best Local Similarity 100.0%; Pred. No. 1.8e-09;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3716 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 546 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 592

RESULT 33
US-09-381-488-6
; Sequence 6, Application US/09381488
; Patent No. 6441135
; GENERAL INFORMATION:
; APPLICANT: Kiefer, Michael C.
; APPLICANT: Gibson, Helen L.
; APPLICANT: Fitzpatrick, Paul A.
; APPLICANT: Barr, Philip J.
; TITLE OF INVENTION: A NOVEL Bak BINDING PROTEIN, DNA
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSER: BROWN, THERESA A.
; STREET: 1560 Broadway, Suite 1200
; CITY: Denver
; STATE: CO
; COUNTRY: USA
; ZIP: 80202
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/09/381.488
; FILING DATE: 11-Feb-2000
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWN, THERESA A.
; REGISTRATION NUMBER: 32,547
; REFERENCE/DOCKET NUMBER: 4147-15-PUS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 975 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
```

```
; NAME/KEY: CDS
; LOCATION: 26..478
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-381-488-6

Query Match 1.2%; Score 47; DB 4; Length 975;
Best Local Similarity 100.0%; Pred. No. 1.7e-09;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3716 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
|
|
|
Db 919 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 965
|
|
|

RESULT 34
US-08-665-716-1
; Sequence 1, Application US/08665716
; Patent No. 5789222
; GENERAL INFORMATION:
; APPLICANT: KELLY, ROSEMARIE
; APPLICANT: REGISTER, ELIZABETH A
; APPLICANT: MASUREKAR, PRAKASH S
; TITLE OF INVENTION: P5C REDUCTASE GENE FROM ZALERION
; TITLE OF INVENTION: ARBORICOLA
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MERCK & CO., INC.
; STREET: 126 E. LINCOLN AVENUE
; CITY: RAHWAY
; STATE: NEW JERSEY
; COUNTRY: US
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/665,716
; FILING DATE: 23-JUN-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: KORSN, ELLIOTT
; REGISTRATION NUMBER: 32,705
; REFERENCE/DOCKET NUMBER: 19453PV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 908-594-5493
; TELEFAX: 908-594-4720
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1147 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 47..960
US-08-665-716-1

Query Match 1.2%; Score 47; DB 1; Length 1147;
Best Local Similarity 100.0%; Pred. No. 1.7e-09;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3716 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
|
|
|
Db 1071 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1117
|
|
|

RESULT 35
US-08-713-000-7
; Sequence 7, Application US/08713000
; Patent No. 5850020
; GENERAL INFORMATION:
; APPLICANT: Blockberg, Leonard N.
; APPLICANT: Havukkala, Ilkka
; APPLICANT: Grierson, Alastair
; TITLE OF INVENTION: MATERIALS AND METHODS FOR THE
; MODIFICATION OF PLANT LIGNIN CONTENT
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Speckman Picard PLLC
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/713,000
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sleath, Janet
; REGISTRATION NUMBER: 37,007
; REFERENCE/DOCKET NUMBER: 11000.1003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-269-0565
; TELEFAX: 206-269-0563
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1454 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-713-000-7

Query Match 1.2%; Score 47; DB 2; Length 1454;
Best Local Similarity 100.0%; Pred. No. 1.6e-09;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3716 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
|
|
|
Db 1405 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1451
|
|
|

RESULT 36
US-08-975-316-7
; Sequence 7, Application US/08975316
; Patent No. 5952486
; GENERAL INFORMATION:
; APPLICANT: BLOCKSBERG, Leonard N., HAVUKKALA, Ilkka
; APPLICANT: and GRIERSON Alastair W.
; TITLE OF INVENTION: MATERIALS AND METHODS FOR
; THE MODIFICATION OF PLANT LIGNIN CONTENT
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Ann W. Speckman
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/975,316
```

```
; FILING DATE: 1997-11-21
; CLASSIFICATION: 800
; PRIOR APPLICATION NUMBER: 08/713,000
; FILING DATE: September 11, 1996
; FILING DATE: September 11, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: SLEATH, Janet
; REGISTRATION NUMBER: 37,007
; REFERENCE/DOCKET NUMBER: 11000/1003C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-269-0565
; TELEFAX: 206-269-0563
; TELEX:
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1454 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-975-316-7

Query Match 1.2%; Score 47; DB 2; Length 1454;
Best Local Similarity 100.0%; Pred. No. 1.6e-09;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3716 ATGAAAAA 3762
Db 1405 ATGAAAAA 1451

RESULT 37
US-09-211-710-7
; Sequence 7, Application US/09211710A
; Patent No. 6204434
; GENERAL INFORMATION:
; APPLICANT: Bloksberg, Leonard N.
; APPLICANT: Havukkala, Ilkka
; APPLICANT: Grierson, Alastair
; TITLE OF INVENTION: Materials and Methods for the
; FILE REFERENCE: 11000.1003c3
; CURRENT APPLICATION NUMBER: US/09/211,710A
; CURRENT FILING DATE: 1998-12-14
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 1454
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-211-710-7

Query Match 1.2%; Score 47; DB 3; Length 1454;
Best Local Similarity 100.0%; Pred. No. 1.6e-09;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3716 ATGAAAAA 3762
Db 1405 ATGAAAAA 1451

RESULT 38
US-09-615-192A-7
; Sequence 7, Application US/09615192A
; Patent No. 6410718
; GENERAL INFORMATION:
; APPLICANT: Bloksberg, Leonard N.
; APPLICANT: Havukkala, Ilkka
; TITLE OF INVENTION: Materials and Methods for the
; FILE REFERENCE: 11000.1003c4U
; CURRENT APPLICATION NUMBER: US/09/615,192A
; CURRENT FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 08/975,316
; PRIOR FILING DATE: 1998-10-09
```

```
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: US 08/713,000
; PRIOR FILING DATE: 1996-09-11
; PRIOR APPLICATION NUMBER: US 09/169,789
; PRIOR FILING DATE: 1998-10-09
; NUMBER OF SEQ ID NOS: 405
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 1454
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-615-192A-7

Query Match 1.2%; Score 47; DB 4; Length 1454;
Best Local Similarity 100.0%; Pred. No. 1.6e-09;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3716 ATGAAAAA 3762
Db 1405 ATGAAAAA 1451

RESULT 39
US-09-169-789-7
; Sequence 7, Application US/09169789
; Patent No. 6653528
; GENERAL INFORMATION:
; APPLICANT: Bloksberg, Leonard N.
; APPLICANT: Havukkala, Ilkka
; TITLE OF INVENTION: Materials and Methods for the
; FILE REFERENCE: 11000.1003c2
; CURRENT APPLICATION NUMBER: US/09/169,789
; CURRENT FILING DATE: 1998-10-09
; EARLIER APPLICATION NUMBER: US 08/975,316
; EARLIER FILING DATE: 1997-11-21
; EARLIER APPLICATION NUMBER: US 08/713,000
; NUMBER OF SEQ ID NOS: 185
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 1454
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-169-789-7

Query Match 1.2%; Score 47; DB 4; Length 1454;
Best Local Similarity 100.0%; Pred. No. 1.6e-09;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3716 ATGAAAAA 3762
Db 1405 ATGAAAAA 1451

RESULT 40
US-09-615-192A-95
; Sequence 95, Application US/09615192A
; Patent No. 6410718
; GENERAL INFORMATION:
; APPLICANT: Bloksberg, Leonard N.
; APPLICANT: Havukkala, Ilkka
; TITLE OF INVENTION: Materials and Methods for the
; FILE REFERENCE: 11000.1003c4U
; CURRENT APPLICATION NUMBER: US/09/615,192A
; CURRENT FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 08/975,316
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: US 08/713,000
; PRIOR FILING DATE: 1996-09-11
; PRIOR APPLICATION NUMBER: US 09/169,789
; PRIOR FILING DATE: 1998-10-09
```

```
; NUMBER OF SEQ ID NOS: 405
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 95
; LENGTH: 1460
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-615-192A-95

Query Match
Best Local Similarity 1.2%; Score 47; DB 4; Length 1460;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3716 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 1411 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1457

RESULT 41
US-09-169-789-95
; Sequence 95, Application US/09169789
; Patent No. 6653528
; GENERAL INFORMATION:
; APPLICANT: Bloksberg, Leonard N.
; APPLICANT: Havukkala, Ilkka
; TITLE OF INVENTION: Materials and Methods for the
; TITLE OF INVENTION: Modification of Plant Lignin Content
; FILE REFERENCE: 11000.1003C2
; CURRENT APPLICATION NUMBER: US/09/169,789
; CURRENT FILING DATE: 1998-10-09
; EARLIER APPLICATION NUMBER: US 08/975,316
; EARLIER FILING DATE: 1997-11-21
; EARLIER APPLICATION NUMBER: US 08/713,000
; EARLIER FILING DATE: 1996-09-11
; NUMBER OF SEQ ID NOS: 185
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 95
; LENGTH: 1460
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-169-789-95

Query Match
Best Local Similarity 1.2%; Score 47; DB 4; Length 1460;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3716 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 1411 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1457

RESULT 42
US-08-975-316-71
; Sequence 71, Application US/08975316
; Patent No. 5952486
; GENERAL INFORMATION:
; APPLICANT: BLOKBERG, Leonard N., HAVUKKALA, Ilkka
; APPLICANT: and GRIERSON, Alastair W.
; TITLE OF INVENTION: MATERIALS AND METHODS FOR
; TITLE OF INVENTION: THE MODIFICATION OF PLANT LIGNIN CONTENT
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Ann W. Speckman
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/975,316
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/713,000
; FILING DATE: September 11, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: SLEATH, Janet
; REGISTRATION NUMBER: 37,007
; REFERENCE/DOCKET NUMBER: 11000/1003C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-269-0565
; TELEFAX: 206-269-0563
; TELEX:
; INFORMATION FOR SEQ ID NO: 71:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1474 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-975-316-71

Query Match
Best Local Similarity 1.2%; Score 47; DB 2; Length 1474;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3716 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 1425 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1471

RESULT 43
US-09-615-192A-71
; Sequence 71, Application US/09615192A
; Patent No. 6410718
; GENERAL INFORMATION:
; APPLICANT: Bloksberg, Leonard N.
; APPLICANT: Havukkala, Ilkka
; TITLE OF INVENTION: Materials and Methods for the
; TITLE OF INVENTION: Modification of Plant Lignin Content
; FILE REFERENCE: 11000.1003C4U
; CURRENT APPLICATION NUMBER: US/09/615,192A
; CURRENT FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 08/975,316
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: US 08/713,000
; PRIOR FILING DATE: 1996-09-11
; PRIOR APPLICATION NUMBER: US 09/169,789
; PRIOR FILING DATE: 1998-10-09
; NUMBER OF SEQ ID NOS: 405
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 71
; LENGTH: 1474
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-615-192A-71

Query Match
Best Local Similarity 1.2%; Score 47; DB 4; Length 1474;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3716 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 1425 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1471

RESULT 44
US-09-615-192A-402
; Sequence 402, Application US/09615192A
; Patent No. 6410718
; GENERAL INFORMATION:
; APPLICANT: Bloksberg, Leonard N.
; APPLICANT: Havukkala, Ilkka
```

```
; TITLE OF INVENTION: Materials and Methods for the
; FILE OF INVENTION: Modification of Plant Lignin Content
; FILE REFERENCE: 11000.1003c4U
; CURRENT APPLICATION NUMBER: US/09/615,192A
; CURRENT FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 08/975,316
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: US 08/713,000
; PRIOR FILING DATE: 1996-09-11
; PRIOR APPLICATION NUMBER: US 09/169,789
; PRIOR FILING DATE: 1998-10-09
; NUMBER OF SEQ ID NOS: 405
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 402
; LENGTH: 1474
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-615-192A-402

Query Match
Best Local Similarity 100.0%; Score 47; DB 4; Length 1474;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3716 ATGAAAAA... 3762
DB 1425 ATGAAAAA... 1471

RESULT 45
US-09-325-932A-202
; Sequence 202, Application US/09325932A
; Patent No. 6451604
; GENERAL INFORMATION:
; APPLICANT: Flinn, Barry
; APPLICANT: Lasham, Annette
; TITLE OF INVENTION: Compositions affecting programmed cell
; FILE OF INVENTION: death and their use in the modification of forestry plant develo
; FILE REFERENCE: 1022
; CURRENT APPLICATION NUMBER: US/09/325,932A
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 202
; LENGTH: 1474
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-325-932A-202

Query Match
Best Local Similarity 100.0%; Score 47; DB 4; Length 1474;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3716 ATGAAAAA... 3762
DB 1425 ATGAAAAA... 1471

RESULT 46
US-09-169-789-71
; Sequence 71, Application US/09169789
; Patent No. 6653528
; GENERAL INFORMATION:
; APPLICANT: Bloksberg, Leonard N.
; APPLICANT: Havukkala, Ilkka
; TITLE OF INVENTION: Materials and Methods for the
; FILE OF INVENTION: Modification of Plant Lignin Content
; FILE REFERENCE: 11000.1003c2
; CURRENT APPLICATION NUMBER: US/09/169,789
; CURRENT FILING DATE: 1998-10-09
; EARLIER APPLICATION NUMBER: US 08/975,316
; EARLIER FILING DATE: 1997-11-21
; EARLIER APPLICATION NUMBER: US 08/713,000
; EARLIER FILING DATE: 1996-09-11
```

```
; NUMBER OF SEQ ID NOS: 185
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 71
; LENGTH: 1474
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-169-789-71

Query Match
Best Local Similarity 100.0%; Score 47; DB 4; Length 1474;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3716 ATGAAAAA... 3762
DB 1425 ATGAAAAA... 1471

RESULT 47
US-09-169-789-184
; Sequence 184, Application US/09169789
; Patent No. 6653528
; GENERAL INFORMATION:
; APPLICANT: Bloksberg, Leonard N.
; APPLICANT: Havukkala, Ilkka
; TITLE OF INVENTION: Materials and Methods for the
; FILE OF INVENTION: Modification of Plant Lignin Content
; FILE REFERENCE: 11000.1003c2
; CURRENT APPLICATION NUMBER: US/09/169,789
; CURRENT FILING DATE: 1998-10-09
; EARLIER APPLICATION NUMBER: US 08/975,316
; EARLIER FILING DATE: 1997-11-21
; EARLIER APPLICATION NUMBER: US 08/713,000
; EARLIER FILING DATE: 1996-09-11
; NUMBER OF SEQ ID NOS: 185
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 184
; LENGTH: 1474
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-169-789-184

Query Match
Best Local Similarity 100.0%; Score 47; DB 4; Length 1474;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3716 ATGAAAAA... 3762
DB 1425 ATGAAAAA... 1471

RESULT 48
US-09-370-807-7
; Sequence 7, Application US/09370807
; Patent No. 6297034
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Falco, S. Carl
; APPLICANT: Rafalski, J. Antoni
; APPLICANT: Sakai, Hajime
; TITLE OF INVENTION: N-End Rule Pathway Enzymes
; FILE REFERENCE: BB-1199
; CURRENT APPLICATION NUMBER: US/09/370,807
; CURRENT FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: 60/096,225
; EARLIER FILING DATE: August 12, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 7
; LENGTH: 2407
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-370-807-7
```



```

; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 102
; TYPE: DNA
; LENGTH: 199
; ORGANISM: Homo sapien
US-09-736-457-102

Query Match 1.2%; Score 46; DB 4; Length 199;
Best Local Similarity 100.0%; Pred. No. 4.9e-09;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3717 TGAAGAAAAA 3762
DB 154 TGAAGAAAAA 199

RESULT 56
US-09-614-124B-102
; Sequence 102, Application US/09614124B
; Patent No. 6630574
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C9
; CURRENT APPLICATION NUMBER: US/09/614,124B
; CURRENT FILING DATE: 2001-07-11
; NUMBER OF SEQ ID NOS: 1668
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 102
; TYPE: DNA
; LENGTH: 199
; ORGANISM: Homo sapien
US-09-614-124B-102

Query Match 1.2%; Score 46; DB 4; Length 199;
Best Local Similarity 100.0%; Pred. No. 4.9e-09;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3717 TGAAGAAAAA 3762
DB 154 TGAAGAAAAA 199

RESULT 57
US-09-671-325-102
; Sequence 102, Application US/09671325
; Patent No. 6667154
; GENERAL INFORMATION:

```

```

; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedwick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, AiJun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: DIAGNOSIS OF LUNG CANCER
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-07-21
; SOFTWARE: Patent.pm
; SEQ ID NO 10543
; LENGTH: 193
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 36..37
; OTHER INFORMATION: n=a, g, c or t
US-09-621-976-10543

Query Match      1.2%; Score 46; DB 4; Length 193;
Best Local Similarity 100.0%; Pred. No. 4.9e-09;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   3717 TGAAGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db   102  TGAAGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 147

RESULT 54
US-09-702-705-102
; Sequence 102, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedwick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: DIAGNOSIS OF LUNG CANCER
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 102
; LENGTH: 199
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-702-705-102

Query Match      1.2%; Score 46; DB 4; Length 199;
Best Local Similarity 100.0%; Pred. No. 4.9e-09;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   3717 TGAAGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db   154  TGAAGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 199

RESULT 55
US-09-736-457-102
; Sequence 102, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:

```

APPLICANT: Wang, Tongtong
APPLICANT: Baugur, Chaitanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Darrick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
APPLICANT: Fan, Liqun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.478C12
CURRENT APPLICATION NUMBER: US/09/671.325
CURRENT FILING DATE: 2000-09-26
NUMBER OF SEQ ID NOS: 1825
SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO 102
LENGTH: 199
TYPE: DNA
ORGANISM: Homo sapien
US-09-671-325-102

Query Match 1.2%; Score 46; DB 4; Length 199;
Best Local Similarity 100.0%; Pred. No. 4.9e-09;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3717 TGAIAA 3762
Db 154 TGAIAA 199

RESULT 58
US-09-589-184-102
Sequence 102, Application US/09589184
Patent No. 6686447
GENERAL INFORMATION:
APPLICANT: Wang, Tongtong
APPLICANT: Baugur, Chaitanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Darrick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.478C8
CURRENT APPLICATION NUMBER: US/09/589,184
CURRENT FILING DATE: 2000-06-05
NUMBER OF SEQ ID NOS: 827
SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO 102
LENGTH: 199
TYPE: DNA
ORGANISM: Homo sapien
US-09-589-184-102

Query Match 1.2%; Score 46; DB 4; Length 199;
Best Local Similarity 100.0%; Pred. No. 4.9e-09;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3717 TGAIAA 3762
Db 154 TGAIAA 199

RESULT 59
US-09-621-976-16536
Sequence 16536, Application US/09621976
Patent No. 6639063
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Jobert, S.

APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: GENSET.054PR2
CURRENT APPLICATION NUMBER: US/09/621,976
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.pm
SEQ ID NO 16536
LENGTH: 213
TYPE: DNA
ORGANISM: Homo sapiens
US-09-621-976-16536

Query Match 1.2%; Score 46; DB 4; Length 213;
Best Local Similarity 100.0%; Pred. No. 4.9e-09;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3717 TGAIAA 3762
Db 156 TGAIAA 201

RESULT 60
US-09-621-976-16317
Sequence 16317, Application US/09621976
Patent No. 6639063
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Jobert, S.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: GENSET.054PR2
CURRENT APPLICATION NUMBER: US/09/621,976
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.pm
SEQ ID NO 16317
LENGTH: 231
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: 231
OTHER INFORMATION: n=a, g, c or t
US-09-621-976-16317

Query Match 1.2%; Score 46; DB 4; Length 231;
Best Local Similarity 100.0%; Pred. No. 4.8e-09;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3717 TGAIAA 3762
Db 160 TGAIAA 205

RESULT 61
US-09-621-976-9455
Sequence 9455, Application US/09621976
Patent No. 6639063
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Jobert, S.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: GENSET.054PR2
CURRENT APPLICATION NUMBER: US/09/621,976
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.pm
SEQ ID NO 9455
LENGTH: 235
TYPE: DNA
ORGANISM: Homo sapiens


```
Query Match      1.2%; Score 46; DB 4; Length 249;
Best Local Similarity 100.0%; Pred.No. 4.8e-09;
Matches 46; Conservative 0; Mismatches 0; Indels 0
```

```
QY    3717 TGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db     190 TGAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 235
```

```

RESULT 67
US-09-621-976-18893
; Sequence 18893, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 18893
; LENGTH: 250
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-18893

```

```

Query Match      1.2%; Score 46; DB 4; Length 250;
Best Local Similarity 100.0%; Pred. No. 4.8e-09;
Matches 46; Conservative 0; Mismatches 0; Indels 0

Qy 3717 TGAIAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 176 TGAIAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 221

```

```

RESULT 68
US-09-621-376-16294
; Sequence 16294, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 16294
; LENGTH: 259
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-376-16294

```

	Query Match	1.2%	Score 46	DB 4	Length 259
	Best Local Similarity	100.0%	Pred. No. 4.8e-09		
	Matches 46	Conservative 0	Mismatches 0	Indels 0	
Qy	3717	TGAAA	AAA	3762	
Db	189	TGAAA	AAA	234	

RESULT 69
US-09-821-976-18648
; Sequence 18648, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:

```

; APPLICANT: Dumas Milne Edwards, J.B.
;
; APPLICANT: Jobert, S.
;
; APPLICANT: Giordano, J.Y.
;
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
;
; FILE REFERENCE: GENSET.054PR2
;
; CURRENT APPLICATION NUMBER: US/09/621,976
;
; CURRENT FILING DATE: 2000-07-21
;
; NUMBER OF SEQ ID NOS: 19335
;
; SOFTWARE: Patent.pm
;
; SEQ ID NO 18648
;
; LENGTH: 282
;
; TYPE: DNA
;
; ORGANISM: Homo sapiens
;
US-09-621-976-18648

```

```

Query Match      1.2%; Score 46; DB 4; Length 282;
Best Local Similarity 100.0%; Pred. No. 4.7e-09;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3717 TGA##### 3762
      |||||#####
Db 175 TGA##### 220
      |||||#####

```

```

RESULT 70
US-09-621-976-16989
; Sequence 16989, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 16989
; LENGTH: 283
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-16989

```

```

Query Match      1.2%; Score 46; DB 4; Length 283;
Best Local Similarity 100.0%; Pred. No. 4.7e-09;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3717 TGAaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3762
Db 214 TGAaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 259

```

```

RESULT 71
US-09-621-976-16581
? Sequence 16581, Application US/09621976
? Patent No. 6639063
? GENERAL INFORMATION:
? APPLICANT: Dumas Milne Edwards, J.B.
? APPLICANT: Jobert, S.
? APPLICANT: Giordano, J.Y.
? TITLE OF INVENTION: ESTs and Encoded Human Proteins.
? FILE REFERENCE: GENSET 054PR2
? CURRENT APPLICATION NUMBER: US/09/621,976
? CURRENT FILING DATE: 2000-07-21
? NUMBER OF SEQ ID NOS: 19335
? SOFTWARE: Patent.fm
? SEQ ID NO 16581
? LENGTH: 356
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: misc feature

```

```

; LOCATION: 230..231
; OTHER INFORMATION: n=a, g, c or t
US-09-621-976-16581

```

Query Match 1.2%; Score 46; DB 4; Length 356;
Best Local Similarity 100.0%; Pred. No. 4.6e-09;
Matches 46; Conservative 0; Mismatches 0; Indels

Qy	3717	TGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	3762
Db	308	TGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	353

RESULT 72

RESOLU. 72
US-09-621-976-9484
; Sequence 9484, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESFs and Encoded Human Proteins.

FILE REFERENCE: GENSET.054PR2
CURRENT APPLICATION NUMBER: US/09/621,976
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 19335

```

; SOFTWARE: Patent.pm
; SEQ ID NO 9484
; LENGTH: 399

```

US-09-621-976-9484

```
Query Match      1.2%; Score 46; DB 4; Length 399;
Best Local Similarity 100.0%; Pred.No. 4.6e-09;
Matches 46; Conservative 0; Mismatches 0; Indels
```

	3717	TGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	3762
Qy			
	349	TGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	394
Dd			

RESULT 73

US-09-220-132-186/c
; Sequence 186, Application US/09220132
; Patent No. 6506607
; GENERAL INFORMATION.

```

; GENERAL INFORMATION:
; APPLICANT: Shyjan, Andrew W.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT
; TITLE OF INVENTION: OF PROSTATE CANCER THERAPIES AND THE DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 07334-074001
; CURRENT APPLICATION NUMBER: US/09/220,132

```

; CURRENT FILING DATE: 1998-12-23
 ; PRIOR APPLICATION NUMBER: US 60/079,303
 ; PRIOR FILING DATE: 1998-03-25
 ; PRIOR APPLICATION NUMBER: US 60/068,821

```

; PRIOR APPLICATION NUMBER: US
; PRIOR FILING DATE: 1997-12-24
; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: FastSeq for Windows

```

```

; SOFWAKE: FASTSEQ LOI WINDOWS VERSION 4.1
;
; SEQ ID NO 186
;
; LENGTH: 495
;
; TYPE: DNA

```

LIFE: DNA
 ORGANISM: HOM
 US-09-220-132-186

```
Query Match      1.2%; Score 46; DB 4; Length 495;
Best Local Similarity 100.0%; Pred.No. 4.5e-09;
Matches 46; Conservative 0; Mismatches 0; Indels
```

QY	3717	TCGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	3762
D5	64	TCGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	19

[illegible]

RESULT 74

US-09-152-060-26
; Sequence 26, Application US/09152060
; Patent No. 6448230
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.

```

1 AFFILIATION: ROSEN ET AL.
2
3 TITLE OF INVENTION: 28 Human Secreted Proteins
4
5 FILE REFERENCE: P2003P1.US
6
7 CURRENT APPLICATION NUMBER: US/09/152,060
8
9 CURRENT FILING DATE: 1998-09-11
10
11 EARLIER APPLICATION NUMBER: PCT/US98/04858
12
13 EARLIER FILING DATE: 1998-03-12
14

```

; EARLIER FILING DATE: 1998-03-12
 ; EARLIER APPLICATION NUMBER: 60/040,762
 ; EARLIER FILING DATE: 1997-03-14
 ; EARLIER APPLICATION NUMBER: 60/040,710

; EARLIER APPLICATION NUMBER: 60/040,111
 ; EARLIER FILING DATE: 1997-03-14
 ; EARLIER APPLICATION NUMBER: 60/050,934
 ; EARLIER FILING DATE: 1997-05-30

; EARLIER FILING DATE: 1997-05-30
 ; EARLIER APPLICATION NUMBER: 60/048,100
 ; EARLIER FILING DATE: 1997-05-30
 ; EARLIER APPLICATION NUMBER: 60/048,357

; EARLIER FILING DATE: 1987-05-30
 ; EARLIER APPLICATION NUMBER: 60/048,189
 ; EARLIER FILING DATE: 1987-05-30
 ; EARLIER APPLICATION NUMBER: 60/048,355

; EARLIER FILING DATE: 1997-05-30
; EARLIER APPLICATION NUMBER: 60/057,765
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/048,870

EARLIER APPLICATION NUMBER: 60/048,970
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/068,368
EARLIER FILING DATE: 1997-12-10

```

; EARLIER FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 20

```

```
; SEQ ID NO 26
;
; LENGTH: 700
; TYPE: DNA
;
```

```

; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
;

```

```

; LOCATION: (81)
; OTHER INFORMATION: n equals a,t,g, or
; FEATURE:

```

```

; NAME/KEY: SITE
; LOCATION: (659)
; OTHER INFORMATION: n equals a,t,g, or

```

```

; FEATURE:
; NAME/KEY: SITE
; LOCATION: (692)

```

```

;
; OTHER INFORMATION: n equals a,t,g, or
;
; FEATURE:
;
; NAME/KEY: SITE
;

```

US-09-152-060-26

Query Match	1.2%;	Score 4
Best Local Similarity	100.0%;	Pred. 1

Matches 46; Conservative 0; Mismatches 3717 TGAAAAAAAAAAAAAAAAAAAAAAAAA

D_b 612 TGA AAAAAAAAAAAAAAAAAAAAAAAAAA

RESULT 75
US-09-902-331B-3

US 02-3119
; Sequence 3, Application US/09902331B
; Patent No. 6660907
; GENERAL INFORMATION:

; GENERAL INFORMATION:
 ; APPLICANT: Hendrick, Carol A
 ; APPLICANT: Hu, Xu
 ; APPLICANT: Lu, Guihua

```

; AFFILIATION: DU, Guitara
; TITLE OF INVENTION: GENES ENCODING SCI

```

```

COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/197,793
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/688,019;
FILING DATE: 19-APR-1991
APPLICATION NUMBER: 08/004,199
FILING DATE: 13-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Deconti, Giulio A., Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: BBI-006CNP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 953 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 64..681
US-08-197-793-1

Query Match 1.2%; Score 46; DB 1; Length 953;
Best Local Similarity 100.0%; Pred. No. 4.2e-09;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3717 TGA
DB 905 TGA
3762
950

RESULT 78
US-08-636-176-1
Sequence 1, Application US/08636176
Patent No. 5846822
GENERAL INFORMATION:
APPLICANT: Meuer, S.
APPLICANT: Schraven, B.
APPLICANT: Schoenhaut, D.
APPLICANT: Ratnofsky, S.
TITLE OF INVENTION: pp32: A Newly Identified CD45-Associated
TITLE OF INVENTION: Protein
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSER: LAHIVE & COCKFIELD
STREET: 60 STATE STREET, SUITE 510
CITY: BOSTON
STATE: MASSACHUSETTS
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/636,176
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:

```

APPLICATION NUMBER: 07/588,019; 08/004,199
FILING DATE: 19-APR-1991; 13-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: DeConti, Giulio A., Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: BBI-006CNCPC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 953 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 64..681
US-08-636-176-1

Query Match 1.2%; Score 46; DB 2; Length 953;
Best Local Similarity 100.0%; Pred. No. 4.2e-09;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3717 TCGAAA 3762
DB 905 TCGAAA 950

RESULT 79
PCT-US95-01618-1
Sequence 1, Application PC/TUS9501618
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: Pp32: A Newly Identified CD45-Associated
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 STATE STREET, SUITE 510
CITY: BOSTON
STATE: MASSACHUSETTS
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/01618
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 08/197,793
ATTORNEY/AGENT INFORMATION:
NAME: DeConti, Giulio A., Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: BBI-006CPCPC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 953 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 64..681

Query Match 1.2%; Score 46; DB 2; Length 953;
Best Local Similarity 100.0%; Pred. No. 4.2e-09;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3717 TCGAAA 3762
DB 905 TCGAAA 950

RESULT 80
US-09-800-729-14
Sequence 14, Application US/09800729
Patent No. 605592
GENERAL INFORMATION:
APPLICANT: Ni et al.
TITLE OF INVENTION: 32 Human secreted proteins
FILE REFERENCE: P2044P1
CURRENT APPLICATION NUMBER: US/09/800,729
CURRENT FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: PCT/US00/25013
PRIOR FILING DATE: 2000-09-22
PRIOR APPLICATION NUMBER: 60/155,709
PRIOR FILING DATE: 1999-09-24
NUMBER OF SEQ ID NOS: 217
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 14
LENGTH: 997
TYPE: DNA
ORGANISM: Homo sapiens
US-09-800-729-14

Query Match 1.2%; Score 46; DB 4; Length 997;
Best Local Similarity 100.0%; Pred. No. 4.2e-09;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3717 TCGAAA 3762
DB 951 TCGAAA 996

RESULT 81
US-09-257-179-21
Sequence 21, Application US/09257179
Patent No. 6410709
GENERAL INFORMATION:
APPLICANT: Ruben et al.
TITLE OF INVENTION: 29 Human Secreted Proteins
FILE REFERENCE: P2015P1
CURRENT APPLICATION NUMBER: US/09/257,179
CURRENT FILING DATE: 1999-02-25
EARLIER APPLICATION NUMBER: PCT/US98/17709
EARLIER FILING DATE: 1998-08-27
EARLIER APPLICATION NUMBER: 60/056,270
EARLIER FILING DATE: 1997-08-29
EARLIER APPLICATION NUMBER: 60/056,271
EARLIER FILING DATE: 1997-08-29
EARLIER APPLICATION NUMBER: 60/056,247
EARLIER FILING DATE: 1997-08-29
EARLIER APPLICATION NUMBER: 60/056,073
EARLIER FILING DATE: 1997-08-29
NUMBER OF SEQ ID NOS: 128
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 21
LENGTH: 1032
TYPE: DNA
ORGANISM: Homo sapiens
US-09-257-179-21

Query Match 1.2%; Score 46; DB 4; Length 1032;
Best Local Similarity 100.0%; Pred. No. 4.2e-09;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3717 TCGAAA 3762
DB 951 TCGAAA 996

RESULT 81
US-09-257-179-21
Sequence 21, Application US/09257179
Patent No. 6410709
GENERAL INFORMATION:
APPLICANT: Ruben et al.
TITLE OF INVENTION: 29 Human Secreted Proteins
FILE REFERENCE: P2015P1
CURRENT APPLICATION NUMBER: US/09/257,179
CURRENT FILING DATE: 1999-02-25
EARLIER APPLICATION NUMBER: PCT/US98/17709
EARLIER FILING DATE: 1998-08-27
EARLIER APPLICATION NUMBER: 60/056,270
EARLIER FILING DATE: 1997-08-29
EARLIER APPLICATION NUMBER: 60/056,271
EARLIER FILING DATE: 1997-08-29
EARLIER APPLICATION NUMBER: 60/056,247
EARLIER FILING DATE: 1997-08-29
EARLIER APPLICATION NUMBER: 60/056,073
EARLIER FILING DATE: 1997-08-29
NUMBER OF SEQ ID NOS: 128
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 21
LENGTH: 1032
TYPE: DNA
ORGANISM: Homo sapiens
US-09-257-179-21

Query Match 1.2%; Score 46; DB 4; Length 1032;
Best Local Similarity 100.0%; Pred. No. 4.2e-09;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3717 TCGAAA 3762
DB 951 TCGAAA 996


```
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: Patentin Ver: 2.0
; SEQ ID NO 47
; LENGTH: 1149
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-800-729-47

Query Match
Best Local Similarity 1.2%; Score 46; DB 4; Length 1149;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3717 TCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
    |||
Db 1063 TCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1108

RESULT 87
US-09-149-476-41
; Sequence 41, Application US/09149476
; Patent No. 6420526
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002P1
; CURRENT APPLICATION NUMBER: US/09/149,476
; CURRENT FILING DATE: 1998-09-08
; EARLIER APPLICATION NUMBER: PCT/US98/04493
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,533
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,584
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,500
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,587
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,492
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,598
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,613
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,582
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,596
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,612
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,632
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,601
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,580
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,568
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,314
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,569
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,311
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,671
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,674
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,669
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,312
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,313
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,672
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,315
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/056,886
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,877
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,889
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,893
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,630
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,878
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,662
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,872
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,882
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,637
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,903
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,888
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,879
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,880
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,894
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,911
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,636
```

RESULT 89
US-09-205-258-25
; Sequence 25, Application US/09205258
; Patent No. 6525174
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: P2007P1
; CURRENT APPLICATION NUMBER: US/09/205,258
; CURRENT FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884

; TITLE OF INVENTION: NOVEL CSATPT PROTEIN MOLECULES AND USES THEREFOR (AS AMENDED)

; FILE REFERENCE: MNI-051DV1

; CURRENT APPLICATION NUMBER: US/09/221,448A

; CURRENT FILING DATE: 1998-12-28

; NUMBER OF SEQ ID NOS: 23

; SOFTWARE: Patent In Ver. 2.0

; SEQ ID NO 1

; LENGTH: 1315

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (249)..(767)

; OTHER INFORMATION: 'n' at position 1315 may be any nucleotide

US-09-221-448A-1

Query Match 1.2%; Score 46; DB 4; Length 1315;

Best Local Similarity 100.0%; Fred. No. 4.1e-03;

Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3717 TCAA 3762

Db 1245 TCAA 1290

RESULT 93

US-09-149-476-208

; Sequence 208, Application US/09149476

; Patent No. 6420526

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: 186 Human Secreted proteins

; FILE REFERENCE: P2002P1

; CURRENT APPLICATION NUMBER: US/09/149,476

; CURRENT FILING DATE: 1998-09-08

; EARLIER APPLICATION NUMBER: PCT/US98/04493

; EARLIER FILING DATE: 1998-03-06

; EARLIER APPLICATION NUMBER: 60/040,162

; EARLIER FILING DATE: 1997-03-07

; EARLIER APPLICATION NUMBER: 60/040,333

; EARLIER FILING DATE: 1997-03-07

; EARLIER APPLICATION NUMBER: 60/038,621

; EARLIER FILING DATE: 1997-03-07

; EARLIER APPLICATION NUMBER: 60/040,626

; EARLIER FILING DATE: 1997-03-07

; EARLIER APPLICATION NUMBER: 60/040,334

; EARLIER FILING DATE: 1997-03-07

; EARLIER APPLICATION NUMBER: 60/040,336

; EARLIER FILING DATE: 1997-03-07

; EARLIER APPLICATION NUMBER: 60/040,163

; EARLIER FILING DATE: 1997-03-07

; EARLIER APPLICATION NUMBER: 60/047,600

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,615

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,597

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,502

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,633

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,583

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,617

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,618

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,503

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,592

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,581

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,581

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,584

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,500

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,587

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,492

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,598

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,613

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,582

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,596

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,612

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,632

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,601

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/043,580

; EARLIER FILING DATE: 1997-04-11

; EARLIER APPLICATION NUMBER: 60/043,568

; EARLIER FILING DATE: 1997-04-11

; EARLIER APPLICATION NUMBER: 60/043,314

; EARLIER FILING DATE: 1997-04-11

; EARLIER APPLICATION NUMBER: 60/043,569

; EARLIER FILING DATE: 1997-04-11

; EARLIER APPLICATION NUMBER: 60/043,311

; EARLIER FILING DATE: 1997-04-11

; EARLIER APPLICATION NUMBER: 60/043,671

; EARLIER FILING DATE: 1997-04-11

; EARLIER APPLICATION NUMBER: 60/043,674

; EARLIER FILING DATE: 1997-04-11

; EARLIER APPLICATION NUMBER: 60/043,669

; EARLIER FILING DATE: 1997-04-11

; EARLIER APPLICATION NUMBER: 60/043,312

; EARLIER FILING DATE: 1997-04-11

; EARLIER APPLICATION NUMBER: 60/043,313

; EARLIER FILING DATE: 1997-04-11

; EARLIER APPLICATION NUMBER: 60/043,672

; EARLIER FILING DATE: 1997-04-11

; EARLIER APPLICATION NUMBER: 60/043,315

; EARLIER FILING DATE: 1997-04-11

; EARLIER APPLICATION NUMBER: 60/048,974

; EARLIER FILING DATE: 1997-08-06

; EARLIER APPLICATION NUMBER: 60/056,886

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,877

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,889

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,893

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,630

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,878

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,662

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,872

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,882

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,637

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,903

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,888

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,879

```

; EARLIER APPLICATION NUMBER: 60/056,884
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/057,669
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/049,610
; EARLIER FILING DATE: 1997-06-13
; EARLIER APPLICATION NUMBER: 60/061,060
; EARLIER FILING DATE: 1997-10-02

Query Match          1.2%; Score 46; DB 4; Length 1378;
Best Local Similarity 100.0%; Pred. No. 4.1e-09;
Matches 46; Conservative 0; Mismatches 0; Indels

Qy 3717 TGA
Db 1293 TGA

RESULT 94
US-09-585-173B-7
; Sequence 7, Application US/09585173B
; Patent No. 6570063
; GENERAL INFORMATION:
; APPLICANT: Butler, Karlene
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Guttridge, Steven
; APPLICANT: Maxwell, Carl
; TITLE OF INVENTION: Magnesium Chelataase
; FILE REFERENCE: BE1370 US NA
; CURRENT APPLICATION NUMBER: US/09/585,173B
; CURRENT FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: US 60/137,461
; PRIOR FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 7
; LENGTH: 1477
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-585-173B-7

Query Match          1.2%; Score 46; DB 4; Length 1477;
Best Local Similarity 100.0%; Pred. No. 4.1e-09;
Matches 46; Conservative 0; Mismatches 0; Indels

Qy 3717 TGA
Db 1417 TGA

RESULT 95
US-09-907-794A-220
; Sequence 220, Application US/09907794A
; Patent No. 6635468
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.

```

```
/ APPLICANT: Filvaroff, Ellen
/ APPLICANT: Fong, Sherman
/ APPLICANT: Gao, Wei-Qiang
/ APPLICANT: Gerber, Hanspeter
/ APPLICANT: Gerritsen, Mary E.
/ APPLICANT: Goddard, A.
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Grimaldi, Christopher J.
/ APPLICANT: Guney, Austin L.
/ APPLICANT: Hillan, Kenneth, Jr.
/ APPLICANT: Kljavin, Ivar J.
/ APPLICANT: Mather, Jennie P.
/ APPLICANT: Pan, James
/ APPLICANT: Paoni, Nicholas F.
/ APPLICANT: Roy, Margaret Ann
/ APPLICANT: Stewart, Timothy A.
/ APPLICANT: Tunas, Daniel
/ APPLICANT: Williams, P. Mickey
/ APPLICANT: Wood, William, Jr.
/ TITLE OF INVENTION: Acids and Transmembrane Polypeptides and Nucleic
/ FILE REFERENCE: 10466-14
/ CURRENT APPLICATION NUMBER: US/09/905,125A
/ PRIOR FILING DATE: 2001-07-12
/ PRIOR APPLICATION NUMBER: PCT/US00/04414
/ PRIOR FILING DATE: 2000-02-22
/ PRIOR APPLICATION NUMBER: US 60/143,048
/ PRIOR FILING DATE: 1999-07-07
/ PRIOR APPLICATION NUMBER: US 60/145,698
/ PRIOR FILING DATE: 1999-07-26
/ PRIOR APPLICATION NUMBER: US 60/146,222
/ PRIOR FILING DATE: 1999-07-28
/ PRIOR APPLICATION NUMBER: PCT/US99/20594
/ PRIOR FILING DATE: 1999-09-08
/ PRIOR APPLICATION NUMBER: PCT/US99/20944
/ PRIOR FILING DATE: 1999-09-13
/ PRIOR APPLICATION NUMBER: PCT/US99/21090
/ PRIOR FILING DATE: 1999-09-15
/ PRIOR APPLICATION NUMBER: PCT/US99/21547
/ PRIOR FILING DATE: 1999-09-15
/ PRIOR APPLICATION NUMBER: PCT/US99/23089
/ PRIOR FILING DATE: 1999-10-05
/ PRIOR APPLICATION NUMBER: PCT/US99/28214
/ PRIOR FILING DATE: 1999-11-29
/ PRIOR APPLICATION NUMBER: PCT/US99/28313
/ PRIOR FILING DATE: 1999-11-30
/ PRIOR APPLICATION NUMBER: PCT/US99/28564
/ PRIOR FILING DATE: 1999-12-02
/ PRIOR APPLICATION NUMBER: PCT/US99/28565
/ PRIOR FILING DATE: 1999-12-02
/ PRIOR APPLICATION NUMBER: PCT/US99/30095
/ PRIOR FILING DATE: 1999-12-16
/ PRIOR APPLICATION NUMBER: PCT/US99/30911
/ PRIOR FILING DATE: 1999-12-20
/ PRIOR APPLICATION NUMBER: PCT/US99/30999
/ PRIOR FILING DATE: 1999-12-20
/ PRIOR APPLICATION NUMBER: PCT/US00/00219
/ PRIOR FILING DATE: 2000-01-05
/ SEQUENCE ID NOS: 423
/ SEQ ID NO 220
/ LENGTH: 1503
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ US-09-905-125A-220
```

Query Match
Best Local Similarity 1.2%; Score 46; DB 4; Length 1503;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps

CY 3717 TGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
DB 1445 TGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1490


```
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-902-775A-220

Query Match          1.2%; Score 46; DB 4; Length 1503;
Best Local Similarity 100.0%; Pred. No. 4e-09;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3717 TGAAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      1445 TGACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1490

RESULT 98
US-09-244-111-7
; Sequence 7, Application US/09244111
; Patent No. 6566498
; GENERAL INFORMATION:
; APPLICANT: Ni, et al.
; TITLE OF INVENTION: Human Serine Protease and Serpin Polypeptides
; FILE REFERENCE: PF391
; CURRENT APPLICATION NUMBER: US/09/244,111
; CURRENT FILING DATE: 1999-02-04
; EARLIER APPLICATION NUMBER: 60/073,961
; EARLIER FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 7
; LENGTH: 1527
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: {67}..(1335)
US-09-244-111-7

Query Match          1.2%; Score 46; DB 4; Length 1527;
Best Local Similarity 100.0%; Pred. No. 4e-09;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3717 TGAAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      1474 TCAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1519

RESULT 99
US-09-820-004-1
; Sequence 1, Application US/09820004
; Patent No. 6649385
; GENERAL INFORMATION:
; APPLICANT: Wei, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001201
; CURRENT APPLICATION NUMBER: US/09/820,004
; CURRENT FILING DATE: 2001-03-29
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1606
; TYPE: DNA
; ORGANISM: Human
US-09-820-004-1

Query Match          1.2%; Score 46; DB 4; Length 1606;
Best Local Similarity 100.0%; Pred. No. 4e-09;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3717 TGAAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      1532 TCAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1577
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 23, 2004, 08:22:52 ; Search time 1469 Seconds

(without alignments)
11546.259 Million cell updates/sec

Title: US-10-005-907-1

Perfect score: 3762

Sequence: 1 gagaacgggtcactgtga.....aaaaaaaaaaaaaaaa 3762

Scoring table: OLIGO_NUC

Gapex 60.0 , Gapext 60.0

Searched: 2907579 seqs, 225431464 residues

Word size : 0

Total number of hits satisfying chosen parameters: 5815158

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 150 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result.	No.	Score	Query Match	Length	DB ID	Description
1	3762	100.0	3762	15	US-10-005-907-1	Sequence 1, Appli
2	50	1.3	261	13	US-10-424-599-141898	Sequence 141898,
3	50	1.3	1702	9	US-09-925-300-395	Sequence 395, App
4	50	1.3	3602	15	US-10-171-311-67	Sequence 67, Appl
5	49	1.3	476	14	US-10-001-843-87	Sequence 87, Appl
6	49	1.3	606	13	US-10-424-599-15023	Sequence 35023, A
7	49	1.3	1057	15	US-10-106-998-741	Sequence 741, App
8	49	1.3	1373	9	US-09-925-301-10	Sequence 10, Appl
9	49	1.3	1392	15	US-10-198-846-9741	Sequence 9741, Ap
10	49	1.3	1417	13	US-10-424-599-11755	Sequence 11755, A
11	49	1.3	1423	13	US-10-424-599-106798	Sequence 106798,
12	49	1.3	1429	15	US-10-007-926A-34	Sequence 34, Appl
13	49	1.3	1998	9	US-09-764-903-19	Sequence 19, Appl
14	49	1.3	1998	15	US-10-091-391-19	Sequence 19, Appl

C	15	49	1.3	2143	15	US-10-264-049-740	Sequence 740, App
	16	49	1.3	2713	15	US-10-436-523-1	Sequence 1, Appl
	17	49	1.3	2713	15	US-10-436-523-31	Sequence 31, Appl
	18	49	1.3	2713	15	US-10-436-523-32	Sequence 32, Appl
	19	49	1.3	2713	15	US-10-436-523-33	Sequence 33, Appl
	20	49	1.3	2713	15	US-10-436-523-34	Sequence 34, Appl
C	21	49	1.3	2944	15	US-10-128-714-1558	Sequence 1558, Ap
	22	49	1.3	3362	15	US-10-128-714-6558	Sequence 6558, Ap
	23	49	1.3	4237	9	US-09-745-763-20	Sequence 20, Appl
	24	49	1.3	4944	15	US-10-128-714-558	Sequence 558, App
C	25	49	1.3	5362	15	US-09-960-352-6036	Sequence 6036, Ap
	26	48	1.3	337	9	US-10-085-783A-16208	Sequence 16208, A
	27	48	1.3	348	13	US-10-242-535A-16208	Sequence 16208, A
	28	48	1.3	348	16	US-10-424-599-47139	Sequence 47139, A
	29	48	1.3	372	13	US-10-424-599-47139	Sequence 47139, A
	30	48	1.3	401	10	US-09-918-995-16319	Sequence 16319, A
	31	48	1.3	462	13	US-10-424-599-54555	Sequence 54555, A
C	32	48	1.3	544	13	US-10-424-599-46783	Sequence 46783, A
	33	48	1.3	619	13	US-10-351-334-58	Sequence 58, Appl
	34	48	1.3	751	13	US-10-424-599-95989	Sequence 95989, A
C	35	48	1.3	865	9	US-09-770-445-579	Sequence 579, App
	36	48	1.3	1837	13	US-10-219-535-39	Sequence 39, Appl
	37	48	1.3	1837	13	US-10-232-230-39	Sequence 39, Appl
	38	48	1.3	1837	13	US-10-232-224-39	Sequence 39, Appl
	39	48	1.3	1837	15	US-10-227-884-39	Sequence 39, Appl
	40	48	1.3	1837	15	US-10-230-163-39	Sequence 39, Appl
	41	48	1.3	1837	15	US-10-230-338-39	Sequence 39, Appl
	42	48	1.3	1837	15	US-10-218-631-39	Sequence 39, Appl
	43	48	1.3	1837	15	US-10-230-414-39	Sequence 39, Appl
	44	48	1.3	1837	15	US-10-216-159A-39	Sequence 39, Appl
	45	48	1.3	1837	15	US-10-218-849-39	Sequence 39, Appl
	46	48	1.3	1837	15	US-10-227-873-39	Sequence 39, Appl
	47	48	1.3	1837	15	US-10-227-883-39	Sequence 39, Appl
	48	48	1.3	1837	15	US-10-219-076-39	Sequence 39, Appl
	49	48	1.3	1837	15	US-10-230-434-39	Sequence 39, Appl
	50	48	1.3	1837	15	US-10-219-003-39	Sequence 39, Appl
	51	48	1.3	1837	15	US-10-219-075-39	Sequence 39, Appl
	52	48	1.3	1837	15	US-10-219-466-39	Sequence 39, Appl
	53	48	1.3	1837	15	US-10-219-479-39	Sequence 39, Appl
	54	48	1.3	1837	15	US-10-219-481-39	Sequence 39, Appl
	55	48	1.3	1837	15	US-10-230-260-39	Sequence 39, Appl
	56	48	1.3	1837	15	US-10-232-231-39	Sequence 39, Appl
	57	48	1.3	1837	15	US-10-232-233-39	Sequence 39, Appl
	58	48	1.3	1837	15	US-10-216-185-39	Sequence 39, Appl
	59	48	1.3	1837	15	US-10-218-956-39	Sequence 39, Appl
	60	48	1.3	1837	15	US-10-219-468-39	Sequence 39, Appl
	61	48	1.3	1837	15	US-10-219-478-39	Sequence 39, Appl
	62	48	1.3	1837	15	US-10-219-536-39	Sequence 39, Appl
	63	48	1.3	1837	15	US-10-219-536-39	Sequence 39, Appl
	64	48	1.3	1837	15	US-10-233-205-39	Sequence 39, Appl
	65	48	1.3	1837	15	US-10-219-072-39	Sequence 39, Appl
	66	48	1.3	1837	15	US-10-219-470-39	Sequence 39, Appl
	67	48	1.3	1837	15	US-10-219-474-39	Sequence 39, Appl
	68	48	1.3	1837	15	US-10-219-524-39	Sequence 39, Appl
	69	48	1.3	1837	15	US-10-219-528-39	Sequence 39, Appl
	70	48	1.3	1837	15	US-10-227-880-39	Sequence 39, Appl
	71	48	1.3	1837	15	US-10-227-881-39	Sequence 39, Appl
	72	48	1.3	1837	15	US-10-227-882-39	Sequence 39, Appl
	73	48	1.3	1837	15	US-10-230-436-39	Sequence 39, Appl
	74	48	1.3	1837	15	US-10-232-223-39	Sequence 39, Appl
	75	48	1.3	1837	15	US-10-232-225-39	Sequence 39, Appl
	76	48	1.3	1837	15	US-10-232-227-39	Sequence 39, Appl
	77	48	1.3	1837	15	US-10-232-229-39	Sequence 39, Appl
	78	48	1.3	1837	15	US-10-232-234-39	Sequence 39, Appl
	79	48	1.3	1837	15	US-10-219-060-39	Sequence 39, Appl
	80	48	1.3	1837	15	US-10-216-160-39	Sequence 39, Appl
	81	48	1.3	1837	15	US-10-216-162-39	Sequence 39, Appl
	82	48	1.3	1837	15	US-10-216-164-39	Sequence 39, Appl
	83	48	1.3	1837	15	US-10-216-167-39	Sequence 39, Appl
	84	48	1.3	1837	15	US-10-216-168-39	Sequence 39, Appl
	85	48	1.3	1837	15	US-10-219-065-39	Sequence 39, Appl
	86	48	1.3	1837	15	US-10-219-071-39	Sequence 39, Appl
	87	48	1.3	1837	15	US-10-219-074-39	Sequence 39, Appl

88 1.3 1837 15 US-10-219-077-39
89 1.3 1837 15 US-10-219-465-39
90 1.3 1837 15 US-10-219-467-39
91 1.3 1837 15 US-10-219-469-39
92 1.3 1837 15 US-10-219-471-39
93 1.3 1837 15 US-10-219-473-39
94 1.3 1837 15 US-10-219-476-39
95 1.3 1837 15 US-10-219-482-39
96 1.3 1837 15 US-10-227-874-39
97 1.3 1837 15 US-10-227-876-39
98 1.3 1837 15 US-10-227-878-39
99 1.3 1837 15 US-10-229-974-39
100 1.3 1837 15 US-10-230-024-39
101 1.3 1837 15 US-10-230-113-39
102 1.3 1837 15 US-10-230-183-39
103 1.3 1837 15 US-10-230-234-39
104 1.3 1837 15 US-10-230-306-39
105 1.3 1837 15 US-10-230-426-39
106 1.3 1837 15 US-10-230-427-39
107 1.3 1837 15 US-10-230-433-39
108 1.3 1837 15 US-10-230-435-39
109 1.3 1837 15 US-10-230-438-39
110 1.3 1837 15 US-10-232-222-39
111 1.3 1837 15 US-10-219-070-39
112 1.3 1837 15 US-10-219-472-39
113 1.3 1837 15 US-10-219-527-39
114 1.3 1837 15 US-10-227-877-39
115 1.3 1837 15 US-10-216-166-39
116 1.3 1837 15 US-10-218-612-39
117 1.3 1837 15 US-10-216-163-39
118 1.3 1837 15 US-10-218-765-39
119 1.3 1837 15 US-10-219-063-39
120 1.3 1837 15 US-10-219-066-39
121 1.3 1837 15 US-10-219-067-39
122 1.3 1837 15 US-10-219-068-39
123 1.3 1837 15 US-10-219-069-39
124 1.3 1837 15 US-10-219-073-39
125 1.3 1837 15 US-10-219-475-39
126 1.3 1837 15 US-10-219-480-39
127 1.3 1837 15 US-10-219-483-39
128 1.3 1837 15 US-10-219-525-39
129 1.3 1837 15 US-10-219-526-39
130 1.3 1837 15 US-10-219-530-39
131 1.3 1837 15 US-10-219-531-39
132 1.3 1837 15 US-10-219-532-39
133 1.3 1837 15 US-10-219-533-39
134 1.3 1837 15 US-10-230-437-39
135 1.3 1837 15 US-10-232-228-39
136 1.3 1837 15 US-10-232-226-39
137 1.3 1837 15 US-10-230-130-39
138 1.3 1837 15 US-10-424-599-12512
139 1.3 1837 15 US-10-219-535-7
140 1.3 1837 15 US-10-232-230-7
141 1.3 1837 15 US-10-232-224-7
142 1.3 1837 15 US-10-227-884-7
143 1.3 1837 15 US-10-230-163-7
144 1.3 1837 15 US-10-230-338-7
145 1.3 1837 15 US-10-218-631-7
146 1.3 1837 15 US-10-230-414-7
147 1.3 1837 15 US-10-216-159A-7
148 1.3 1837 15 US-10-218-849-7
149 1.3 1837 15 US-10-227-873-7
150 1.3 1837 15 US-10-227-883-7

ALIGNMENTS

RESULT 1
US-10-005-907-1
; Sequence 1, Application US/10005907
; Publication No. US2003016681A1
; GENERAL INFORMATION:

; APPLICANT: Union Chimique Belge, S.A.
; APPLICANT: No. US2003016681A1, Karl
; APPLICANT: Pirozzi, Gregory
; APPLICANT: Einstein, Richard
; TITLE OF INVENTION: NOVEL GENES ASSOCIATED WITH ALLERGIC HYPERSENSITIVITY AND MAST CELL
; TITLE OF INVENTION: ACTIVATION
; FILE REFERENCE: 053529-5005
; CURRENT APPLICATION NUMBER: US/10/005,907
; CURRENT FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 1
; LENGTH: 3762
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (25)..(432)
; OTHER INFORMATION:
US-10-005-907-1

Query Match 100.0%; Score 3762; DB 15; Length 3762;
Best Local Similarity 100.0%; Pred No. 0;
Matches 3762; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGAAACCGAGTCACCTGTGAAAGATGGGAAATTATCTCTCGGAAAACCTCAGTTGCCTG 60
Db 1 GAGAAACCGAGTCACCTGTGAAAGATGGGAAATTATCTCTCGGAAAACCTCAGTTGCCTG 60

Qy 61 GGAGAGATCAAAAGAGCCCGAG 120
Db 61 GGAGAGATCAAAAGAGCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120

Qy 121 ATGACTACATTTTGAAGAAGAACTTCAAGATCAAGATCAAGATCAAGATCAAGATCAAGATCA 180
Db 121 ATGACTACATTTTGAAGAAGAACTTCAAGATCAAGATCAAGATCAAGATCAAGATCA 180

Qy 181 ACTTCTAATCAGAGAAACGAGATGCGAGTGGTCTGAGAGAGTGGTCTACACTGCTCAT 240
Db 181 ACTTCTAATCAGAGAAACGAGATGCGAGTGGTCTGAGAGAGTGGTCTACACTGCTCAT 240

Qy 241 AATCACATCCCCCATCAGAGATCCTCCCTGAGCTCCCAATGATGATGGCTATGAGAACAT 300
Db 241 AATCACATCCCCCATCAGAGATCCTCCCTGAGCTCCCAATGATGATGGCTATGAGAACAT 300

Qy 301 GACTCCCTCAAGAGAAAGTGAGACAGATTTAGAGAAAGTTCAGAGACAGATATGCCCTT 360
Db 301 GACTCCCTCAAGAGAAAGTGAGACAGATTTAGAGAAAGTTCAGAGACAGATATGCCCTT 360

Qy 361 CTTAGGACTTCTGTAGTAGGCTTGTCTGTCACCCATGAGCATGATGATGAGTTGTG 420
Db 361 CTTAGGACTTCTGTAGTAGGCTTGTCTGTCACCCATGAGCATGATGATGAGTTGTG 420

Qy 421 TTTCACACTAAATCTCAGTGGCTTTATACCTTCAGCAATGAGCAATGAGCAATGAGCAAT 480
Db 421 TTTCACACTAAATCTCAGTGGCTTTATACCTTCAGCAATGAGCAATGAGCAATGAGCAAT 480

Qy 481 TAGCAGACTCTGCGAGAGTTGTTTCACTGAGCAGTGCATGAAACATCTCTTCTGGCTA 540
Db 481 TAGCAGACTCTGCGAGAGTTGTTTCACTGAGCAGTGCATGAAACATCTCTTCTGGCTA 540

Qy 541 AAGTTTGAAGAATATTCTTATTATATATCTTATTATATCTTATTATATCTTATTATATCT 600
Db 541 AAGTTTGAAGAATATTCTTATTATATCTTATTATATCTTATTATATCTTATTATATCT 600

Qy 601 TGGCTTAGGTGAATCATAGAAATTCAGCAATGAGCAATGAGCAATGAGCAATGAGCAATG 660
Db 601 TGGCTTAGGTGAATCATAGAAATTCAGCAATGAGCAATGAGCAATGAGCAATGAGCAATG 660

Qy 661 TTGTAAAGTTTGAAGACATGAGGTGATATAAAAAAACTTTCTTTAGGACAAATATGAAAA 720
Db 661 TTGTAAAGTTTGAAGACATGAGGTGATATAAAAAAACTTTCTTTAGGACAAATATGAAAA 720

APPLICANT: Hoersht, Sebastian
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
TITLE OF INVENTION: OF CERVICAL CANCER
FILE REFERENCE: MRI-035
CURRENT APPLICATION NUMBER: US/10/171.311
CURRENT FILING DATE: 2002-06-12
PRIOR APPLICATION NUMBER: US 60/298,159
PRIOR FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/298,155
PRIOR FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/335,936
PRIOR FILING DATE: 2001-11-14
NUMBER OF SEQ ID NOS: 238
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 67
LENGTH: 3602
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: 2087, 2093, 2098
OTHER INFORMATION: n = A,T,C or G
US-10-171-311-67

Query Match 1.3%; Score 50; DB 15; Length 3602;
Best Local Similarity 100.0%; Pred. No. 4e-12;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAAAA 3762
DB 3501 TAAATGAAAAA 3550

RESULT 5
US-10-001-843-87/c
Sequence 87, Application US/10001843
Publication No. US20020132255A1
GENERAL INFORMATION:
APPLICANT: Salceda, Susana
APPLICANT: Macina, Roberto
APPLICANT: Recipon, Herve
APPLICANT: Caferkey, Robert
APPLICANT: Sun, Yongming
APPLICANT: Liu, Chenghua
APPLICANT: Turner, Leah
TITLE OF INVENTION: Compositions and Methods Relating to Breast Specific Genes and P
FILE REFERENCE: DEX-0267
CURRENT APPLICATION NUMBER: US/10/001.843
CURRENT FILING DATE: 2001-11-20
PRIOR APPLICATION NUMBER: 60/249,992
PRIOR FILING DATE: 2000-11-20
NUMBER OF SEQ ID NOS: 218
SOFTWARE: PatentIn version 3.1
SEQ ID NO 87
LENGTH: 476
TYPE: DNA
ORGANISM: Homo sapien
US-10-001-843-87

Query Match 1.3%; Score 49; DB 14; Length 476;
Best Local Similarity 100.0%; Pred. No. 1e-11;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA 3762
DB 126 AAATGAAAAA 78

RESULT 6
US-10-424-599-35023
Sequence 35023, Application US/10424599
Publication No. US20040031072A1

GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 35023
LENGTH: 606
TYPE: DNA
ORGANISM: Glycine max
FEATURE:
NAME/KEY: unsure
LOCATION: (1)-(606)
OTHER INFORMATION: unsure at all n locations
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MFT3847_131628C.1
US-10-424-599-35023

Query Match 1.3%; Score 49; DB 13; Length 606;
Best Local Similarity 100.0%; Pred. No. 1e-11;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA 3762
DB 85 AAATGAAAAA 133

RESULT 7
US-10-106-698-741
Sequence 741, Application US/10106698
Publication No. US20030109690A1
GENERAL INFORMATION:
APPLICANT: Ruben et al.
TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptides
FILE REFERENCE: PA005P1
CURRENT APPLICATION NUMBER: US/10/106,698
CURRENT FILING DATE: 2002-03-27
PRIOR APPLICATION NUMBER: PCT/US00/26524
PRIOR FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: US 60/157,137
PRIOR FILING DATE: 1999-09-29
PRIOR APPLICATION NUMBER: US 60/163,280
PRIOR FILING DATE: 1999-11-03
NUMBER OF SEQ ID NOS: 8564
SOFTWARE: PatentIn Ver. 3.0
SEQ ID NO 741
LENGTH: 1057
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (1056)-(1057)
OTHER INFORMATION: n equals a,t,g, or c
US-10-106-698-741

Query Match 1.3%; Score 49; DB 15; Length 1057;
Best Local Similarity 100.0%; Pred. No. 1e-11;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA 3762
DB 1006 AAATGAAAAA 1054

RESULT 8
US-09-925-301-10
Sequence 10, Application US/09925301
Patent No. US20020052308A1


```
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: US/09/925,301
; PRIOR FILING DATE: 2001-08-10
; PRIOR FILING DATE: 2000-03-08
; PRIOR FILING DATE: 2000-03-08
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 10
; LENGTH: 1373
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1364)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (1373)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-301-10

Query Match      1.3%; Score 49; DB 9; Length 1373;
Best Local Similarity 100.0%; Pred. No. 1.1e-11;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 1294 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1342

RESULT 9
US-10-198-846-9741/c
; Sequence 9741, Application US/10198846
; Publication No. US20030099741
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9741
; LENGTH: 1392
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-198-846-9741

Query Match      1.3%; Score 49; DB 15; Length 1392;
Best Local Similarity 100.0%; Pred. No. 1.1e-11;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 374 AAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 326

RESULT 10
US-10-424-599-11755
; Sequence 11755, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 11755
; LENGTH: 1417
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_110622C.1
US-10-424-599-11755

Query Match      1.3%; Score 49; DB 13; Length 1417;
Best Local Similarity 100.0%; Pred. No. 1.1e-11;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 1354 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1402

RESULT 11
US-10-424-599-106798
; Sequence 106798, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 106798
; LENGTH: 1423
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_67454C.1
US-10-424-599-106798

Query Match      1.3%; Score 49; DB 13; Length 1423;
Best Local Similarity 100.0%; Pred. No. 1.1e-11;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 1339 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1387

RESULT 12
US-10-007-926A-34
; Sequence 34, Application US/10007926A
; Publication No. US20030143539A1
; GENERAL INFORMATION:
; APPLICANT: BERTUCCI, FRANCOIS
; APPLICANT: HOULGATTE, REMI
; APPLICANT: BIRNBAUM, DANIEL
; APPLICANT: NGUYEN, CATHERINE
; APPLICANT: VIENS, PATRICE
; APPLICANT: FERT, VINCENT
; TITLE OF INVENTION: GENE EXPRESSION PROFILING OF PRIMARY BREAST CARCINOMAS
; TITLE OF INVENTION: USING ARRAYS OF CANDIDATE GENES
; FILE REFERENCE: 1546-R-00
```

;	CURRENT APPLICATION NUMBER: US/10/007,926A	;	PRIOR FILING DATE: 2000-08-14	;	PRIOR FILING DATE: 2000-08-14	;	PRIOR APPLICATION NUMBER: 60/220,963	;
;	CURRENT FILING DATE: 2001-12-07	;	PRIOR FILING DATE: 2000-07-26	;	PRIOR FILING DATE: 2000-07-26	;	PRIOR APPLICATION NUMBER: 60/217,496	;
;	PRIOR APPLICATION NUMBER: 60/254,090	;	PRIOR FILING DATE: 2000-12-08	;	PRIOR FILING DATE: 2000-07-11	;	PRIOR APPLICATION NUMBER: 60/225,447	;
;	NUMBER OF SEQ ID NOS: 468	;	SOFTWARE: PatentIn Ver. 2.1	;	PRIOR FILING DATE: 2000-08-14	;	PRIOR APPLICATION NUMBER: 60/218,290	;
;	SEQ ID NO 34	;	LENGTH: 1429	;	PRIOR FILING DATE: 2000-07-14	;	PRIOR APPLICATION NUMBER: 60/225,757	;
;	TYPE: DNA	;	ORGANISM: Homo sapiens	;	PRIOR FILING DATE: 2000-08-14	;	PRIOR APPLICATION NUMBER: 60/226,868	;
;	FEATURE:	;	OTHER INFORMATION: protein phosphatase 4 (formerly x),	;	PRIOR FILING DATE: 2000-08-22	;	PRIOR APPLICATION NUMBER: 60/216,647	;
;	US-10-007-926A-34	;	US-10-007-926A-34	;	PRIOR FILING DATE: 2000-07-07	;	PRIOR APPLICATION NUMBER: 60/225,267	;
;	Query Match	;	1.3%; Score 49; DB 15; Length 1429;	;	PRIOR FILING DATE: 2000-08-14	;	PRIOR APPLICATION NUMBER: 60/216,880	;
;	Best Local Similarity	;	100.0%; Pred. No. 1.1e-11;	;	PRIOR FILING DATE: 2000-07-07	;	PRIOR APPLICATION NUMBER: 60/225,270	;
;	Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	;		;	PRIOR FILING DATE: 2000-08-14	;	PRIOR APPLICATION NUMBER: 60/251,869	;
;		;		;	PRIOR FILING DATE: 2000-12-08	;	PRIOR APPLICATION NUMBER: 60/235,834	;
;		;		;	PRIOR FILING DATE: 2000-09-27	;	PRIOR APPLICATION NUMBER: 60/234,274	;
;		;		;	PRIOR FILING DATE: 2000-09-21	;	PRIOR APPLICATION NUMBER: 60/234,223	;
;		;		;	PRIOR FILING DATE: 2000-09-21	;	PRIOR APPLICATION NUMBER: 60/228,924	;
;		;		;	PRIOR FILING DATE: 2000-08-30	;	PRIOR APPLICATION NUMBER: 60/224,518	;
;		;		;	PRIOR FILING DATE: 2000-08-14	;	PRIOR APPLICATION NUMBER: 60/236,369	;
;		;		;	PRIOR FILING DATE: 2000-09-29	;	PRIOR APPLICATION NUMBER: 60/224,519	;
;		;		;	PRIOR FILING DATE: 2000-08-14	;	PRIOR APPLICATION NUMBER: 60/220,964	;
;		;		;	PRIOR FILING DATE: 2000-07-26	;	PRIOR APPLICATION NUMBER: 60/241,809	;
;		;		;	PRIOR FILING DATE: 2000-10-20	;	PRIOR APPLICATION NUMBER: 60/249,299	;
;		;		;	PRIOR FILING DATE: 2000-11-17	;	PRIOR APPLICATION NUMBER: 60/236,327	;
;		;		;	PRIOR FILING DATE: 2000-09-29	;	PRIOR APPLICATION NUMBER: 60/241,785	;
;		;		;	PRIOR FILING DATE: 2000-10-20	;	PRIOR APPLICATION NUMBER: 60/244,617	;
;		;		;	PRIOR FILING DATE: 2000-11-01	;	PRIOR APPLICATION NUMBER: 60/225,268	;
;		;		;	PRIOR FILING DATE: 2000-08-14	;	PRIOR APPLICATION NUMBER: 60/236,368	;
;		;		;	PRIOR FILING DATE: 2000-09-29	;	PRIOR APPLICATION NUMBER: 60/251,856	;
;		;		;	PRIOR FILING DATE: 2000-12-08	;	PRIOR APPLICATION NUMBER: 60/251,868	;
;		;		;	PRIOR FILING DATE: 2000-12-08	;	PRIOR APPLICATION NUMBER: 60/229,344	;
;		;		;	PRIOR FILING DATE: 2000-09-01	;	PRIOR APPLICATION NUMBER: 60/234,997	;
;		;		;	PRIOR FILING DATE: 2000-09-25	;	PRIOR APPLICATION NUMBER: 60/229,343	;
;		;		;	PRIOR FILING DATE: 2000-09-01	;	PRIOR APPLICATION NUMBER: 60/229,345	;
;		;		;	PRIOR FILING DATE: 2000-09-01	;	PRIOR APPLICATION NUMBER: 60/229,287	;
;		;		;	PRIOR FILING DATE: 2000-09-01	;	PRIOR APPLICATION NUMBER: 60/229,513	;
;		;		;	PRIOR FILING DATE: 2000-09-05	;	PRIOR APPLICATION NUMBER: 60/231,413	;
;		;		;	PRIOR FILING DATE: 2000-09-08	;	PRIOR APPLICATION NUMBER: 60/229,509	;
;		;		;	PRIOR FILING DATE: 2000-09-05	;		;

; PRIOR APPLICATION NUMBER: 60/236,367
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/237,039
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,038
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/236,370
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/236,802
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,037
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,040
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/240,960
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/239,935
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/239,937
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/241,787
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,474
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/246,532
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/249,216
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,210
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/226,681
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,759
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/225,213
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/227,182
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,214
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/235,836
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/230,438
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/215,135
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 60/225,266
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/249,218
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,208
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,213
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,212
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,207
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,245
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,244
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,217
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,211
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,215
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,264
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,214
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,297

; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/232,400
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/231,242
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/232,081
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/232,080
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,414
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,244
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/233,064
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/233,063
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,397
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,399
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,401
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/241,808
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,826
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,786
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,221
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,475
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/231,243
; PRIOR FILING DATE: 2000-09-08

Query Match 1.3%; Score 49; DB 15; Length 1998;

Best Local Similarity 100.0%; Pred. No. 1.1e-11; Mismatches 0; Indels 0; Gaps 0;

Matches 49; Conservative 0;

Qy 3714 AAATGAAAAA 3762
Db 1906 AAATGAAAAA 1954

RESULT 15

US-10-264-049-740/c
; Sequence 740, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA133P1
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 740
; LENGTH: 2143
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (10)..(10)
; OTHER INFORMATION: n equals a,t,g, or c

OY		3714	A AATGAAAAAAAAAAAAAAAAAAAAA	3762
DQ		2650	A AATGAAAAAAAAAAAAAAAAAAAAA	2698

: APPLICANT: Fraser, Christopher C.
 : TITLE OF INVENTION: CD2000 AND CD2001 MOLECULES, AND USES THEREOF
 : FILE REFERENCE: 7853-244-999
 : CURRENT APPLICATION NUMBER: US/10/436,523
 : CURRENT FILING DATE: 2003-05-12
 : PRIOR APPLICATION NUMBER: US/10/007,303

```
/ PRIOR FILING DATE: 2001-11-20
/ PRIOR APPLICATION NUMBER: 09/706,167
/ PRIOR FILING DATE: 2000-11-03
/ NUMBER OF SEQ ID NOS: 100
/ SOFTWARE: Patent in version 3.1
/ SEQ ID NO 34
/ LENGTH: 2713
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-436-523-34

Query Match      1.3%; Score 49; DB 15; Length 2713;
Best Local Similarity 100.0%; Pred. No. 1.1e-11;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA... 3762
Db 2650 AAATGAAAAA... 2698

RESULT 21
US-10-128-714-1558/c
/ Sequence 1558, Application US/10128714
/ Publication No. US20030119013A1
/ GENERAL INFORMATION:
/ APPLICANT: Jiang, Bo
/ APPLICANT: Hu, Wengqi
/ APPLICANT: Tishkoff, Daniel
/ APPLICANT: Zamudio, Carlos
/ APPLICANT: Eroshkin, Alexey M
/ APPLICANT: Lemieux, Sebastien M
/ TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
/ FILE REFERENCE: 10182-018-999
/ CURRENT APPLICATION NUMBER: US/10/128,714
/ PRIOR FILING DATE: 2002-04-23
/ PRIOR APPLICATION NUMBER: US 60/285,697
/ PRIOR FILING DATE: 2001-04-23
/ PRIOR APPLICATION NUMBER: US 60/287,066
/ PRIOR FILING DATE: 2001-04-27
/ PRIOR APPLICATION NUMBER: US 60/295,890
/ PRIOR FILING DATE: 2001-06-05
/ PRIOR APPLICATION NUMBER: US 60/303,899
/ PRIOR FILING DATE: 2001-07-09
/ PRIOR APPLICATION NUMBER: US 60/316,362
/ PRIOR FILING DATE: 2001-08-31
/ NUMBER OF SEQ ID NOS: 8603
/ SOFTWARE: Patent in version 3.1
/ SEQ ID NO 1558
/ LENGTH: 2944
/ TYPE: DNA
/ ORGANISM: Aspergillus fumigatus
US-10-128-714-1558

Query Match      1.3%; Score 49; DB 15; Length 2944;
Best Local Similarity 100.0%; Pred. No. 1.1e-11;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA... 3762
Db 941 AAATGAAAAA... 893

RESULT 22
US-10-128-714-6558/c
/ Sequence 6558, Application US/10128714
/ Publication No. US20030119013A1
/ GENERAL INFORMATION:
/ APPLICANT: Jiang, Bo
/ APPLICANT: Hu, Wengqi
/ APPLICANT: Tishkoff, Daniel
/ APPLICANT: Zamudio, Carlos
/ APPLICANT: Eroshkin, Alexey M
/ TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
/ FILE REFERENCE: 10182-018-999
/ CURRENT APPLICATION NUMBER: US/10/128,714
/ PRIOR FILING DATE: 2002-04-23
/ PRIOR APPLICATION NUMBER: US 60/285,697
/ PRIOR FILING DATE: 2001-04-23
/ PRIOR APPLICATION NUMBER: US 60/287,066
/ PRIOR FILING DATE: 2001-04-27
/ PRIOR APPLICATION NUMBER: US 60/295,890
/ PRIOR FILING DATE: 2001-06-05
/ PRIOR APPLICATION NUMBER: US 60/303,899
/ PRIOR FILING DATE: 2001-07-09
/ PRIOR APPLICATION NUMBER: US 60/316,362
/ PRIOR FILING DATE: 2001-08-31
/ NUMBER OF SEQ ID NOS: 8603
/ SOFTWARE: Patent in version 3.1
/ SEQ ID NO 1558
/ LENGTH: 2944
/ TYPE: DNA
/ ORGANISM: Aspergillus fumigatus
US-10-128-714-1558

APPLICANT: Lemieux, Sebastien M
/ TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
/ FILE REFERENCE: 10182-018-999
/ CURRENT APPLICATION NUMBER: US/10/128,714
/ PRIOR FILING DATE: 2002-04-23
/ PRIOR APPLICATION NUMBER: US 60/285,697
/ PRIOR FILING DATE: 2001-04-23
/ PRIOR APPLICATION NUMBER: US 60/287,066
/ PRIOR FILING DATE: 2001-04-27
/ PRIOR APPLICATION NUMBER: US 60/295,890
/ PRIOR FILING DATE: 2001-06-05
/ PRIOR APPLICATION NUMBER: US 60/303,899
/ PRIOR FILING DATE: 2001-07-09
/ PRIOR APPLICATION NUMBER: US 60/316,362
/ PRIOR FILING DATE: 2001-08-31
/ NUMBER OF SEQ ID NOS: 8603
/ SOFTWARE: Patent in version 3.1
/ SEQ ID NO 6558
/ LENGTH: 3362
/ TYPE: DNA
/ ORGANISM: Aspergillus fumigatus
US-10-128-714-6558

Query Match      1.3%; Score 49; DB 15; Length 3362;
Best Local Similarity 100.0%; Pred. No. 1.1e-11;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA... 3762
Db 1359 AAATGAAAAA... 1311

RESULT 23
US-09-745-763-20
/ Sequence 20, Application US/09745763
/ Patent No. US20020065394A1
/ GENERAL INFORMATION:
/ APPLICANT: Jacobs, Kenneth
/ APPLICANT: McCoy, John M.
/ APPLICANT: Lavallic, Edward R.
/ APPLICANT: Collins-Racie, Lisa A.
/ APPLICANT: Evans, Cheryl
/ APPLICANT: Merberg, David
/ APPLICANT: Treacy, Maurice
/ APPLICANT: Spaulding, Vikki
/ TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
/ NUMBER OF SEQUENCES: 219
/ CORRESPONDENCE ADDRESS:
/ ADDRESS: Genetics Institute, Inc.
/ STREET: 87 CambridgePark Drive
/ CITY: Cambridge
/ STATE: MA
/ COUNTRY: U.S.A.
/ ZIP: 02140
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent in Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/745,763
/ FILING DATE: 18-Jun-2000
/ CLASSIFICATION: <Unknown>
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Sprunger, Suzanne A.
/ REGISTRATION NUMBER: 41,323
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (617) 498-8284
/ TELEFAX: (617) 876-5851
/ INFORMATION FOR SEQ ID NO: 20:
/ SEQUENCE CHARACTERISTICS:
/
```

```
;
; LENGTH: 4237 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-09-745-763-20

Query Match
  1.3%; Score 49; DB 9; Length 4237;
Best Local Similarity 100.0%; Pred. No. 1.1e-11;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AATGAAAAA
Db 2422 AATGAAAAA

RESULT 24
US-10-714-558/c
; Sequence 558, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wenqi
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Ershkin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
; FILE REFERENCE: 10182-018-999
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 558
; LENGTH: 4944
; TYPE: DNA
; ORGANISM: Aspergillus fumigatus
US-10-128-714-558

Query Match
  1.3%; Score 49; DB 15; Length 4944;
Best Local Similarity 100.0%; Pred. No. 1.2e-11;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AATGAAAAA
Db 1941 AATGAAAAA

RESULT 25
US-10-714-5558/c
; Sequence 5558, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wenqi
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Ershkin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
; FILE REFERENCE: 10182-018-999
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 558
; LENGTH: 4944
; TYPE: DNA
; ORGANISM: Aspergillus fumigatus
US-10-128-714-558

Query Match
  1.3%; Score 49; DB 15; Length 4944;
Best Local Similarity 100.0%; Pred. No. 1.2e-11;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AATGAAAAA
Db 1941 AATGAAAAA

RESULT 26
US-09-960-352-6036
; Sequence 6036, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT FILING DATE: 2001-09-24
; CURRENT APPLICATION NUMBER: US/09/960,352
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 6036
; LENGTH: 337
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 26-LIB3058-015-Q1-X1-G5
US-09-960-352-6036

Query Match
  1.3%; Score 48; DB 9; Length 337;
Best Local Similarity 100.0%; Pred. No. 2.9e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAAAAA
Db 138 AATGAAAAA

RESULT 27
US-10-085-783A-16208
; Sequence 16208, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liaw, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
```

```
;
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10/128,714
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 5558
; LENGTH: 5362
; TYPE: DNA
; ORGANISM: Aspergillus fumigatus
US-10-128-714-5558

Query Match
  1.3%; Score 49; DB 15; Length 5362;
Best Local Similarity 100.0%; Pred. No. 1.2e-11;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AATGAAAAA
Db 2359 AATGAAAAA

RESULT 26
US-09-960-352-6036
; Sequence 6036, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT FILING DATE: 2001-09-24
; CURRENT APPLICATION NUMBER: US/09/960,352
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 6036
; LENGTH: 337
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 26-LIB3058-015-Q1-X1-G5
US-09-960-352-6036

Query Match
  1.3%; Score 48; DB 9; Length 337;
Best Local Similarity 100.0%; Pred. No. 2.9e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAAAAA
Db 138 AATGAAAAA

RESULT 27
US-10-085-783A-16208
; Sequence 16208, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liaw, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
```

; SEQ ID NO 47139	
; LENGTH: 372	
; TYPE: DNA	
; ORGANISM: Glycine max	
; FEATURE:	
; OTHER INFORMATION: Clone ID: PAT_MRT3847_142572C.1	
US-10-424-599-47139	
Query Match	
Best Local Similarity 1.3%; Score 48; DB 13; Length 372;	
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	3715 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
DB	58 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 105
RESULT 30	
US-09-918-995-16319	
; Sequence 16319, Application US/09918995	
; Publication No. US20030073623A1	
; GENERAL INFORMATION:	
; APPLICANT: Hyseq, Inc.	
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED	
; FILE REFERENCE: 20411-756	
; CURRENT APPLICATION NUMBER: US/09/918,995	
; PRIOR FILING DATE: 2001-07-30	
; PRIOR APPLICATION NUMBER: US/09/235,076	
; PRIOR FILING DATE: 1999-01-20	
; NUMBER OF SEQ ID NOS: 38054	
; SOFTWARE: FastSeq for Windows Version 3.0	
; SEQ ID NO 16319	
; LENGTH: 401	
; TYPE: DNA	
; ORGANISM: Homo sapiens	
US-09-918-995-16319	
Query Match	
Best Local Similarity 1.3%; Score 48; DB 10; Length 401;	
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	3715 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
DB	84 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 131
RESULT 31	
US-10-424-599-54555	
; Sequence 54555, Application US/10424599	
; Publication No. US20040031072A1	
; GENERAL INFORMATION:	
; APPLICANT: La Rosa Thomas J	
; APPLICANT: Kovalic David K	
; APPLICANT: Zhou Yihua	
; APPLICANT: Cao Yongwei	
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated	
; FILE REFERENCE: 38-21(53223)B	
; CURRENT APPLICATION NUMBER: US/10/424,599	
; PRIOR FILING DATE: 2003-04-28	
; NUMBER OF SEQ ID NOS: 285684	
; SEQ ID NO 54555	
; LENGTH: 462	
; TYPE: DNA	
; ORGANISM: Glycine max	
; FEATURE:	
; OTHER INFORMATION: Clone ID: PAT_MRT3847_20276C.1	
US-10-424-599-54555	
Query Match	
Best Local Similarity 1.3%; Score 48; DB 13; Length 462;	
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
; PRIOR FILING DATE: 2001-07-13	
; PRIOR APPLICATION NUMBER: US 60/275,017	
; PRIOR FILING DATE: 2001-03-12	
; PRIOR APPLICATION NUMBER: US 60/271,955	
; PRIOR FILING DATE: 2001-02-28	
; NUMBER OF SEQ ID NOS: 58994	
; SOFTWARE: PatentIn version 3.2	
; SEQ ID NO 16208	
; LENGTH: 348	
; TYPE: DNA	
; ORGANISM: Human	
US-10-085-783A-16208	
Query Match	
Best Local Similarity 1.3%; Score 48; DB 13; Length 348;	
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	3715 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
DB	271 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 318
RESULT 28	
US-10-242-535A-16208	
; Sequence 16208, Application US/10242535A	
; Publication No. US20040013663A1	
; GENERAL INFORMATION:	
; APPLICANT: ChondroGene Inc.	
; APPLICANT: Liew, C.C.	
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis	
; FILE REFERENCE: 4231/2005	
; CURRENT APPLICATION NUMBER: US/10/242,535A	
; CURRENT FILING DATE: 2002-09-12	
; PRIOR APPLICATION NUMBER: US 10/085,783	
; PRIOR FILING DATE: 2002-02-28	
; PRIOR APPLICATION NUMBER: US 60/305,340	
; PRIOR FILING DATE: 2001-07-13	
; PRIOR APPLICATION NUMBER: US 60/275,017	
; PRIOR FILING DATE: 2001-03-12	
; PRIOR APPLICATION NUMBER: US 60/271,955	
; PRIOR FILING DATE: 2001-02-28	
; NUMBER OF SEQ ID NOS: 58994	
; SOFTWARE: PatentIn version 3.2	
; SEQ ID NO 16208	
; LENGTH: 348	
; TYPE: DNA	
; ORGANISM: Human	
US-10-242-535A-16208	
Query Match	
Best Local Similarity 1.3%; Score 48; DB 16; Length 348;	
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	3715 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
DB	271 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 318
RESULT 29	
US-10-424-599-47139	
; Sequence 47139, Application US/10424599	
; Publication No. US20040031072A1	
; GENERAL INFORMATION:	
; APPLICANT: La Rosa Thomas J	
; APPLICANT: Kovalic David K	
; APPLICANT: Zhou Yihua	
; APPLICANT: Cao Yongwei	
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With	
; FILE REFERENCE: 38-21(53223)B	
; CURRENT APPLICATION NUMBER: US/10/424,599	
; PRIOR FILING DATE: 2003-04-28	
; NUMBER OF SEQ ID NOS: 285684	
; SEQ ID NO 47139	
; LENGTH: 372	
; TYPE: DNA	
; ORGANISM: Glycine max	
; FEATURE:	
; OTHER INFORMATION: Clone ID: PAT_MRT3847_142572C.1	
US-10-424-599-47139	
Query Match	
Best Local Similarity 1.3%; Score 48; DB 13; Length 372;	
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	3715 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
DB	58 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 105
RESULT 30	
US-09-918-995-16319	
; Sequence 16319, Application US/09918995	
; Publication No. US20030073623A1	
; GENERAL INFORMATION:	
; APPLICANT: Hyseq, Inc.	
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED	
; FILE REFERENCE: 20411-756	
; CURRENT APPLICATION NUMBER: US/09/918,995	
; PRIOR FILING DATE: 2001-07-30	
; PRIOR APPLICATION NUMBER: US/09/235,076	

QY 3715 AATGAAAAA 1.3%; Score 48; DB 13; Length 544;
Best Local Similarity 100.0%; Pred. No. 3e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 120 AATGAAAAA 1.3%; Score 48; DB 13; Length 544;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 32

US-10-424-599-46783/c
; Sequence 46783, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 46783
; LENGTH: 544
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE: Glycine max
; OTHER INFORMATION: Clone ID: PAT_MRT3847_142247C.1
US-10-424-599-46783

Query Match 1.3%; Score 48; DB 13; Length 544;
Best Local Similarity 100.0%; Pred. No. 3e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAAAAA 1.3%; Score 48; DB 13; Length 544;
Best Local Similarity 100.0%; Pred. No. 3e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 57 AATGAAAAA 1.3%; Score 48; DB 13; Length 544;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 33

US-10-351-334-58
; Sequence 58, Application US/10351334
; Publication No. US20040034196A1
; GENERAL INFORMATION:
; APPLICANT: Komatsoulis et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: P2031P2
; CURRENT APPLICATION NUMBER: US/10/351,334
; CURRENT FILING DATE: 2003-01-27
; PRIOR APPLICATION NUMBER: 60/350,898
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/489,847
; PRIOR FILING DATE: 2000-01-24
; PRIOR APPLICATION NUMBER: PCT/US99/17130
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: 60/094,657
; PRIOR FILING DATE: 1998-07-30
; PRIOR APPLICATION NUMBER: 60/095,486
; PRIOR FILING DATE: 1998-08-05
; PRIOR APPLICATION NUMBER: 60/096,319
; PRIOR FILING DATE: 1998-08-12
; PRIOR APPLICATION NUMBER: 60/095,454
; PRIOR FILING DATE: 1998-08-06
; PRIOR APPLICATION NUMBER: 60/095,455
; PRIOR FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 58
; LENGTH: 619
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (526)

; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (619)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-351-334-58

Query Match 1.3%; Score 48; DB 13; Length 619;
Best Local Similarity 100.0%; Pred. No. 3e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3715 AATGAAAAA 1.3%; Score 48; DB 13; Length 619;
Best Local Similarity 100.0%; Pred. No. 3e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 566 AATGAAAAA 1.3%; Score 48; DB 13; Length 619;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 34

US-10-424-599-95989
; Sequence 95989, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 95989
; LENGTH: 751
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)-(751)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_5768C.1
US-10-424-599-95989

Query Match 1.3%; Score 48; DB 13; Length 751;
Best Local Similarity 100.0%; Pred. No. 3e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3715 AATGAAAAA 1.3%; Score 48; DB 13; Length 751;
Best Local Similarity 100.0%; Pred. No. 3e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 513 AATGAAAAA 1.3%; Score 48; DB 13; Length 751;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 35

US-09-770-445-579/c
; Sequence 579, Application US/09770445
; Patent No. US20020023281A1
; GENERAL INFORMATION:
; APPLICANT: Gorlach, Jorn
; APPLICANT: An, Yong-Oiang
; APPLICANT: Hamilton, Carol M.
; APPLICANT: Price, Jennifer L.
; APPLICANT: Raines, Tracy M.
; APPLICANT: Yu, Yang
; APPLICANT: Rameaka, Joshua G.
; APPLICANT: Page, Amy
; APPLICANT: Matthew, Abraham V.
; APPLICANT: Ledford, Brooke L.
; APPLICANT: Woessner, Jeffrey P.
; APPLICANT: Haas, William David
; APPLICANT: Garcia, Carlos A.
; APPLICANT: Krieger, Naja
; APPLICANT: Slader, Ted
; APPLICANT: Davis, Keith R.

RESULT 38
US-10-232-224-39
: Sequence 39. Application US/10232224

;; PRIOR APPLICATION NUMBER: 60/099811
;; PRIOR FILING DATE: 1998-09-10
;; PRIOR APPLICATION NUMBER: 60/099812
;; PRIOR FILING DATE: 1998-09-10
;; PRIOR APPLICATION NUMBER: 60/099816
;; PRIOR FILING DATE: 1998-09-10
;; PRIOR APPLICATION NUMBER: 60/100038
;; PRIOR FILING DATE: 1998-09-11
;; PRIOR APPLICATION NUMBER: 60/100385
;; PRIOR FILING DATE: 1998-09-15
;; PRIOR APPLICATION NUMBER: 60/100390
;; PRIOR FILING DATE: 1998-09-15
;; PRIOR APPLICATION NUMBER: 60/100627
;; PRIOR FILING DATE: 1998-09-16
;; PRIOR APPLICATION NUMBER: 60/100848
;; PRIOR FILING DATE: 1998-09-18
;; PRIOR APPLICATION NUMBER: 60/100919
;; PRIOR FILING DATE: 1998-09-17
;; PRIOR APPLICATION NUMBER: 60/101477
;; PRIOR FILING DATE: 1998-09-23
;; PRIOR APPLICATION NUMBER: 60/101738
;; PRIOR FILING DATE: 1998-09-24
;; PRIOR APPLICATION NUMBER: 60/101741
;; PRIOR FILING DATE: 1998-09-24
;; PRIOR APPLICATION NUMBER: 60/101786
;; PRIOR FILING DATE: 1998-09-25
;; PRIOR APPLICATION NUMBER: 60/101916
;; PRIOR FILING DATE: 1998-09-24
;; PRIOR APPLICATION NUMBER: 60/101922
;; PRIOR FILING DATE: 1998-09-24
;; PRIOR APPLICATION NUMBER: 60/106178
;; PRIOR FILING DATE: 1998-10-28
;; PRIOR APPLICATION NUMBER: 60/106248
;; PRIOR FILING DATE: 1998-10-29
;; PRIOR APPLICATION NUMBER: 60/106464
;; PRIOR FILING DATE: 1998-10-30
;; PRIOR APPLICATION NUMBER: 60/106905
;; PRIOR FILING DATE: 1998-11-03
;; PRIOR APPLICATION NUMBER: 60/108787
;; PRIOR FILING DATE: 1998-11-17
;; PRIOR APPLICATION NUMBER: 60/108801
;; PRIOR FILING DATE: 1998-11-17
;; PRIOR APPLICATION NUMBER: 60/108849
;; PRIOR FILING DATE: 1998-11-18
;; PRIOR APPLICATION NUMBER: 60/112422
;; PRIOR FILING DATE: 1998-12-15
;; PRIOR APPLICATION NUMBER: 60/113296
;; PRIOR FILING DATE: 1998-12-22
;; PRIOR APPLICATION NUMBER: 60/113605
;; PRIOR FILING DATE: 1998-12-23
;; PRIOR APPLICATION NUMBER: 60/113621
;; PRIOR FILING DATE: 1998-12-23
;; PRIOR APPLICATION NUMBER: 60/115558
;; PRIOR FILING DATE: 1999-01-12
;; PRIOR APPLICATION NUMBER: 60/115565
;; PRIOR FILING DATE: 1999-01-12
;; PRIOR APPLICATION NUMBER: 60/115733
;; PRIOR FILING DATE: 1999-01-12
;; PRIOR APPLICATION NUMBER: 60/119549
;; PRIOR FILING DATE: 1999-02-10
;; PRIOR APPLICATION NUMBER: 60/123618
;; PRIOR FILING DATE: 1999-03-10
;; PRIOR APPLICATION NUMBER: 60/125259
;; PRIOR FILING DATE: 1999-03-19
;; PRIOR APPLICATION NUMBER: 60/125775
;; PRIOR FILING DATE: 1999-03-23
;; PRIOR APPLICATION NUMBER: 60/126773
;; PRIOR FILING DATE: 1999-03-29
;; PRIOR APPLICATION NUMBER: 60/127887
;; PRIOR FILING DATE: 1999-04-05
;; PRIOR APPLICATION NUMBER: 60/130232
;; PRIOR FILING DATE: 1999-04-21
;; PRIOR APPLICATION NUMBER: 60/131022

;; PRIOR FILING DATE: 1999-04-26
;; PRIOR APPLICATION NUMBER: 60/131270
;; PRIOR FILING DATE: 1999-04-27
;; PRIOR APPLICATION NUMBER: 60/131291
;; PRIOR FILING DATE: 1999-04-27
;; PRIOR APPLICATION NUMBER: 60/131445
;; PRIOR FILING DATE: 1999-04-28
;; PRIOR APPLICATION NUMBER: 60/134287
;; PRIOR FILING DATE: 1999-05-14
;; PRIOR APPLICATION NUMBER: 60/140650
;; PRIOR FILING DATE: 1999-06-22
;; PRIOR APPLICATION NUMBER: 60/140723
;; PRIOR FILING DATE: 1999-06-22
;; PRIOR APPLICATION NUMBER: 60/141037
;; PRIOR FILING DATE: 1999-06-23
;; PRIOR APPLICATION NUMBER: 60/144758
;; PRIOR FILING DATE: 1999-07-20
;; PRIOR APPLICATION NUMBER: 60/145698
;; PRIOR FILING DATE: 1999-07-26
;; PRIOR APPLICATION NUMBER: 60/146222
;; PRIOR FILING DATE: 1999-07-28
;; PRIOR APPLICATION NUMBER: 60/146963
;; PRIOR FILING DATE: 1999-08-03
;; PRIOR APPLICATION NUMBER: 60/149320
;; PRIOR FILING DATE: 1999-08-17
;; PRIOR APPLICATION NUMBER: 60/149638
;; PRIOR FILING DATE: 1999-08-17
;; PRIOR APPLICATION NUMBER: 60/151733
;; PRIOR FILING DATE: 1999-08-31
;; PRIOR APPLICATION NUMBER: 60/164418
;; PRIOR FILING DATE: 1999-11-09
;; PRIOR APPLICATION NUMBER: 60/166361
;; PRIOR FILING DATE: 1999-11-16
;; PRIOR APPLICATION NUMBER: 60/169445
;; PRIOR FILING DATE: 1999-12-07
;; PRIOR APPLICATION NUMBER: 60/169495
;; PRIOR FILING DATE: 1999-12-07
;; PRIOR APPLICATION NUMBER: 60/169835

Query Match 1.3%; Score 48; DB 15; Length 1837;

Best Local Similarity 100.0%; Pred. No. 3.2e-11;

Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3715 AATGAA 3762
|||||
Db 1729 AATGAA 1776

RESULT 40

US-10-230-163-39
; Sequence 39, Application US/10230163
; Publication No. US20030036635A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3530PIC96
; CURRENT APPLICATION NUMBER: US/10/230,163
; CURRENT FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: 10/119,480
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17

PRIOR APPLICATION NUMBER: 60/062287
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/063549
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/064103
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/069873
PRIOR FILING DATE: 1997-12-17
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/084441
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/085323
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/086392
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089905
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090691
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090695
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/095302
PRIOR FILING DATE: 1998-08-04
PRIOR APPLICATION NUMBER: 60/095318
PRIOR FILING DATE: 1998-08-04
PRIOR APPLICATION NUMBER: 60/095916
PRIOR FILING DATE: 1998-08-10
PRIOR APPLICATION NUMBER: 60/096146
PRIOR FILING DATE: 1998-08-11
PRIOR APPLICATION NUMBER: 60/096791
PRIOR FILING DATE: 1998-08-17
PRIOR APPLICATION NUMBER: 60/097986
PRIOR FILING DATE: 1998-08-26
PRIOR APPLICATION NUMBER: 60/098544
PRIOR FILING DATE: 1998-08-31
PRIOR APPLICATION NUMBER: 60/099596
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: 60/099598
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: 60/099803
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099811
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099812
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099816
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/100038
PRIOR FILING DATE: 1998-09-11
PRIOR APPLICATION NUMBER: 60/100385
PRIOR FILING DATE: 1998-09-15
PRIOR APPLICATION NUMBER: 60/100390
PRIOR FILING DATE: 1998-09-15
PRIOR APPLICATION NUMBER: 60/100627
PRIOR FILING DATE: 1998-09-16
PRIOR APPLICATION NUMBER: 60/100848
PRIOR FILING DATE: 1998-09-18
PRIOR APPLICATION NUMBER: 60/100919
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: 60/101477
PRIOR FILING DATE: 1998-09-23
PRIOR APPLICATION NUMBER: 60/101738
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/101741
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/101786
PRIOR FILING DATE: 1998-09-25
PRIOR APPLICATION NUMBER: 60/101916
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/101922
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/106178
PRIOR FILING DATE: 1998-10-28
PRIOR APPLICATION NUMBER: 60/106248
PRIOR FILING DATE: 1998-10-29
PRIOR APPLICATION NUMBER: 60/106464
PRIOR FILING DATE: 1998-10-30
PRIOR APPLICATION NUMBER: 60/106905
PRIOR FILING DATE: 1998-11-03
PRIOR APPLICATION NUMBER: 60/108787
PRIOR FILING DATE: 1998-11-17
PRIOR APPLICATION NUMBER: 60/108801
PRIOR FILING DATE: 1998-11-17
PRIOR APPLICATION NUMBER: 60/108849
PRIOR FILING DATE: 1998-11-18
PRIOR APPLICATION NUMBER: 60/112422
PRIOR FILING DATE: 1998-12-15
PRIOR APPLICATION NUMBER: 60/113296
PRIOR FILING DATE: 1998-12-22
PRIOR APPLICATION NUMBER: 60/113605
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/113621
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/115558
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/115565
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/115733
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/119549
PRIOR FILING DATE: 1999-02-10
PRIOR APPLICATION NUMBER: 60/123618
PRIOR FILING DATE: 1999-03-10
PRIOR APPLICATION NUMBER: 60/125259
PRIOR FILING DATE: 1999-03-19
PRIOR APPLICATION NUMBER: 60/125775
PRIOR FILING DATE: 1999-03-23
PRIOR APPLICATION NUMBER: 60/126773
PRIOR FILING DATE: 1999-03-29
PRIOR APPLICATION NUMBER: 60/127887
PRIOR FILING DATE: 1999-04-05
PRIOR APPLICATION NUMBER: 60/130232
PRIOR FILING DATE: 1999-04-21
PRIOR APPLICATION NUMBER: 60/131022
PRIOR FILING DATE: 1999-04-26
PRIOR APPLICATION NUMBER: 60/131270
PRIOR FILING DATE: 1999-04-27
PRIOR APPLICATION NUMBER: 60/131291
PRIOR FILING DATE: 1999-04-27
PRIOR APPLICATION NUMBER: 60/131445
PRIOR FILING DATE: 1999-04-28

```
; PRIOR APPLICATION NUMBER: 60/134287
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/140650
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: 60/140723
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: 60/141037
; PRIOR FILING DATE: 1999-06-23
; PRIOR APPLICATION NUMBER: 60/144758
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 60/145698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: 60/146222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: 60/145963
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: 60/149320
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: 60/149638
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: 60/151733
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: 60/164418
; PRIOR FILING DATE: 1999-11-09
; PRIOR APPLICATION NUMBER: 60/166361
; PRIOR FILING DATE: 1999-11-16
; PRIOR APPLICATION NUMBER: 60/169445
; PRIOR FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: 60/169495
; PRIOR FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: 60/169835
```

```
Query Match 1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 3715 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
DB 1729 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1776
```

RESULT 41

```
US-10-230-338-39
; Sequence 39, Application US/10230338
; Publication No. US20030044934A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3530PIC92
; CURRENT APPLICATION NUMBER: US/10/230,338
; PRIOR FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: 10/119,480
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 39
; LENGTH: 1837
; TYPE: DNA
; ORGANISM: Homo Sapien
```

```
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 39
; LENGTH: 1837
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-230-338-39
```

```
Query Match 1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 3715 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
DB 1729 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1776
```

RESULT 42

```
US-10-218-631-39
; Sequence 39, Application US/10218631
; Publication No. US20030045897A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3530PIC14
; CURRENT APPLICATION NUMBER: US/10/218,631
; PRIOR FILING DATE: 2002-08-12
; PRIOR APPLICATION NUMBER: 10/119,480
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 39
; LENGTH: 1837
; TYPE: DNA
; ORGANISM: Homo Sapien
```

US-10-218-631-39

Query Match 1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3715 AATGAAAAA 3762
Db 1729 AATGAAAAA 1776

RESULT 43

US-10-230-414-39
; Sequence 39, Application US/10220414
; Publication No. US20030050448A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3530F1C98
; CURRENT APPLICATION NUMBER: US/10/230,414
; CURRENT FILING DATE: 2002-08-28
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 39
; LENGTH: 1837
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-230-414-39

US-10-230-414-39

US-10-230-414-39

US-10-230-414-39

US-10-230-414-39

US-10-230-414-39

US-10-230-414-39

US-10-230-414-39

US-10-230-414-39

US-10-230-414-39

US-10-230-414-39

US-10-230-414-39

US-10-230-414-39

US-10-230-414-39

US-10-230-414-39

US-10-230-414-39

US-10-230-414-39

US-10-230-414-39

US-10-230-414-39

US-10-230-414-39

US-10-230-414-39

US-10-230-414-39

US-10-230-414-39

US-10-230-414-39

US-10-230-414-39

US-10-230-414-39

US-10-230-414-39

US-10-230-414-39

US-10-230-414-39

US-10-230-414-39

US-10-230-414-39

US-10-230-414-39

US-10-230-414-39

US-10-230-414-39

US-10-230-414-39

US-10-230-414-39

US-10-230-414-39

US-10-230-414-39

US-10-230-414-39

US-10-230-414-39

US-10-230-414-39

US-10-230-414-39

US-10-230-414-39

US-10-230-414-39

US-10-230-414-39

US-10-230-414-39

US-10-230-414-39

US-10-230-414-39

US-10-230-414-39

US-10-230-414-39


```
; CURRENT APPLICATION NUMBER: US/10/218,849
; CURRENT FILING DATE: 2002-08-12
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 39
; LENGTH: 1837
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-218-849-39

Query Match          1.3%  Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAAAAA 3762
    |||||
DB 1729 AATGAAAAA 1776

RESULT 46
US-10-227-873-39
; Sequence 39, Application US/10227873
; Publication No. US20030073816A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3530PLC72
; CURRENT APPLICATION NUMBER: US/10/227,873
; CURRENT FILING DATE: 2002-08-26
; PRIOR APPLICATION NUMBER: 10/119,480
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/081819
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081955
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082804
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/086392
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089905
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090691
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/095302
; PRIOR FILING DATE: 1998-08-04
; PRIOR APPLICATION NUMBER: 60/095318
; PRIOR FILING DATE: 1998-08-04
; PRIOR APPLICATION NUMBER: 60/095916
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: 60/096146
; PRIOR FILING DATE: 1998-08-11
; PRIOR APPLICATION NUMBER: 60/096791
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 60/097986
; PRIOR FILING DATE: 1998-08-26
; PRIOR APPLICATION NUMBER: 60/098544
; PRIOR FILING DATE: 1998-08-31
; PRIOR APPLICATION NUMBER: 60/099596
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099598
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099803
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099811
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099812
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099816
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/100038
; PRIOR FILING DATE: 1998-09-11
; PRIOR APPLICATION NUMBER: 60/100385
; PRIOR FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: 60/100390
; PRIOR FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: 60/100627
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: 60/100848
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/100919
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: 60/101477
; PRIOR FILING DATE: 1998-09-23
; PRIOR APPLICATION NUMBER: 60/101738
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/101741
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/101786
; PRIOR FILING DATE: 1998-09-25
; PRIOR APPLICATION NUMBER: 60/101916
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/101922
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/106178
; PRIOR FILING DATE: 1998-10-28
; PRIOR APPLICATION NUMBER: 60/106248
; PRIOR FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: 60/106464
; PRIOR FILING DATE: 1998-10-30
```

PRIOR APPLICATION NUMBER: 60/106905
PRIOR FILING DATE: 1998-11-03
PRIOR APPLICATION NUMBER: 60/108787
PRIOR FILING DATE: 1998-11-17
PRIOR APPLICATION NUMBER: 60/108801
PRIOR FILING DATE: 1998-11-17
PRIOR APPLICATION NUMBER: 60/108849
PRIOR FILING DATE: 1998-11-18
PRIOR APPLICATION NUMBER: 60/112422
PRIOR FILING DATE: 1998-12-15
PRIOR APPLICATION NUMBER: 60/113296
PRIOR FILING DATE: 1998-12-22
PRIOR APPLICATION NUMBER: 60/113605
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/113621
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/115558
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/115565
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/115733
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/119549
PRIOR FILING DATE: 1999-02-10
PRIOR APPLICATION NUMBER: 60/123618
PRIOR FILING DATE: 1999-03-10
PRIOR APPLICATION NUMBER: 60/125259
PRIOR FILING DATE: 1999-03-19
PRIOR APPLICATION NUMBER: 60/125775
PRIOR FILING DATE: 1999-03-23
PRIOR APPLICATION NUMBER: 60/126773
PRIOR FILING DATE: 1999-03-29
PRIOR APPLICATION NUMBER: 60/127887
PRIOR FILING DATE: 1999-04-05
PRIOR APPLICATION NUMBER: 60/130232
PRIOR FILING DATE: 1999-04-21
PRIOR APPLICATION NUMBER: 60/131022
PRIOR FILING DATE: 1999-04-26
PRIOR APPLICATION NUMBER: 60/131270
PRIOR FILING DATE: 1999-04-27
PRIOR APPLICATION NUMBER: 60/131291
PRIOR FILING DATE: 1999-04-27
PRIOR APPLICATION NUMBER: 60/131445
PRIOR FILING DATE: 1999-04-28
PRIOR APPLICATION NUMBER: 60/134287
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: 60/140650
PRIOR FILING DATE: 1999-06-22
PRIOR APPLICATION NUMBER: 60/140723
PRIOR FILING DATE: 1999-06-22
PRIOR APPLICATION NUMBER: 60/141037
PRIOR FILING DATE: 1999-06-23
PRIOR APPLICATION NUMBER: 60/144758
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/145698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: 60/146222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: 60/146963
PRIOR FILING DATE: 1999-08-03
PRIOR APPLICATION NUMBER: 60/149320
PRIOR FILING DATE: 1999-08-17
PRIOR APPLICATION NUMBER: 60/149638
PRIOR FILING DATE: 1999-08-17
PRIOR APPLICATION NUMBER: 60/151733
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: 60/164418
PRIOR FILING DATE: 1999-11-09
PRIOR APPLICATION NUMBER: 60/166361
PRIOR FILING DATE: 1999-11-16
PRIOR APPLICATION NUMBER: 60/169445
PRIOR FILING DATE: 1999-12-07
PRIOR APPLICATION NUMBER: 60/169495

PRIOR FILING DATE: 1999-12-07
PRIOR APPLICATION NUMBER: 60/169835
Query Match 1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred.No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3715 AATGAA 3762
DB 1729 AATGAA 1776
RESULT 47
US-10-227-883-39
Sequence 39, Application US/10227883
Publication NO. US20030073817A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Desnoyers, Luc
APPLICANT: Gerritsen, Mary
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Smith, Victoria
APPLICANT: Stephan, Jean-Philippe F.
APPLICANT: Watanabe, Colin L.
APPLICANT: Wood, William I.
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3530PIC78
CURRENT APPLICATION NUMBER: US/10/227,883
CURRENT FILING DATE: 2002-08-26
PRIOR APPLICATION NUMBER: 10/119,480
PRIOR FILING DATE: 2002-04-09
PRIOR APPLICATION NUMBER: 60/059113
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/062387
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/063549
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/064103
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/069873
PRIOR FILING DATE: 1997-12-17
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/084441
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/085323
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/086392
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089905
PRIOR FILING DATE: 1998-06-18

1 PRIOR APPLICATION NUMBER: 60/090472
2 PRIOR FILING DATE: 1998-06-24
3 PRIOR APPLICATION NUMBER: 60/090557
4 PRIOR FILING DATE: 1998-06-24
5 PRIOR APPLICATION NUMBER: 60/090691
6 PRIOR FILING DATE: 1998-06-25
7 PRIOR APPLICATION NUMBER: 60/090695
8 PRIOR FILING DATE: 1998-06-25
9 PRIOR APPLICATION NUMBER: 60/091982
10 PRIOR FILING DATE: 1998-07-07
11 PRIOR APPLICATION NUMBER: 60/095302
12 PRIOR FILING DATE: 1998-08-04
13 PRIOR APPLICATION NUMBER: 60/095318
14 PRIOR FILING DATE: 1998-08-04
15 PRIOR APPLICATION NUMBER: 60/095916
16 PRIOR FILING DATE: 1998-08-10
17 PRIOR APPLICATION NUMBER: 60/096146
18 PRIOR FILING DATE: 1998-08-11
19 PRIOR APPLICATION NUMBER: 60/096791
20 PRIOR FILING DATE: 1998-08-17
21 PRIOR APPLICATION NUMBER: 60/097986
22 PRIOR FILING DATE: 1998-08-26
23 PRIOR APPLICATION NUMBER: 60/098544
24 PRIOR FILING DATE: 1998-08-31
25 PRIOR APPLICATION NUMBER: 60/099596
26 PRIOR FILING DATE: 1998-09-09
27 PRIOR APPLICATION NUMBER: 60/099598
28 PRIOR FILING DATE: 1998-09-09
29 PRIOR APPLICATION NUMBER: 60/099803
30 PRIOR FILING DATE: 1998-09-10
31 PRIOR APPLICATION NUMBER: 60/099811
32 PRIOR FILING DATE: 1998-09-10
33 PRIOR APPLICATION NUMBER: 60/099812
34 PRIOR FILING DATE: 1998-09-10
35 PRIOR APPLICATION NUMBER: 60/099816
36 PRIOR FILING DATE: 1998-09-10
37 PRIOR APPLICATION NUMBER: 60/100038
38 PRIOR FILING DATE: 1998-09-11
39 PRIOR APPLICATION NUMBER: 60/100385
40 PRIOR FILING DATE: 1998-09-15
41 PRIOR APPLICATION NUMBER: 60/100390
42 PRIOR FILING DATE: 1998-09-15
43 PRIOR APPLICATION NUMBER: 60/100627
44 PRIOR FILING DATE: 1998-09-16
45 PRIOR APPLICATION NUMBER: 60/100848
46 PRIOR FILING DATE: 1998-09-18
47 PRIOR APPLICATION NUMBER: 60/100919
48 PRIOR FILING DATE: 1998-09-17
49 PRIOR APPLICATION NUMBER: 60/101477
50 PRIOR FILING DATE: 1998-09-23
51 PRIOR APPLICATION NUMBER: 60/101738
52 PRIOR FILING DATE: 1998-09-24
53 PRIOR APPLICATION NUMBER: 60/101741
54 PRIOR FILING DATE: 1998-09-24
55 PRIOR APPLICATION NUMBER: 60/101786
56 PRIOR FILING DATE: 1998-09-25
57 PRIOR APPLICATION NUMBER: 60/101916
58 PRIOR FILING DATE: 1998-09-24
59 PRIOR APPLICATION NUMBER: 60/101922
60 PRIOR FILING DATE: 1998-09-24
61 PRIOR APPLICATION NUMBER: 60/106178
62 PRIOR FILING DATE: 1998-10-28
63 PRIOR APPLICATION NUMBER: 60/106248
64 PRIOR FILING DATE: 1998-10-29
65 PRIOR APPLICATION NUMBER: 60/106464
66 PRIOR FILING DATE: 1998-10-30
67 PRIOR APPLICATION NUMBER: 60/106905
68 PRIOR FILING DATE: 1998-11-03
69 PRIOR APPLICATION NUMBER: 60/108787
70 PRIOR FILING DATE: 1998-11-17
71 PRIOR APPLICATION NUMBER: 60/108801
72 PRIOR FILING DATE: 1998-11-17
73 PRIOR APPLICATION NUMBER: 60/108849

74 PRIOR FILING DATE: 1998-11-18
75 PRIOR APPLICATION NUMBER: 60/112422
76 PRIOR FILING DATE: 1998-12-15
77 PRIOR APPLICATION NUMBER: 60/113296
78 PRIOR FILING DATE: 1998-12-22
79 PRIOR APPLICATION NUMBER: 60/113605
80 PRIOR FILING DATE: 1998-12-23
81 PRIOR APPLICATION NUMBER: 60/113621
82 PRIOR FILING DATE: 1998-12-23
83 PRIOR APPLICATION NUMBER: 60/115558
84 PRIOR FILING DATE: 1999-01-12
85 PRIOR APPLICATION NUMBER: 60/115565
86 PRIOR FILING DATE: 1999-01-12
87 PRIOR APPLICATION NUMBER: 60/115733
88 PRIOR FILING DATE: 1999-01-12
89 PRIOR APPLICATION NUMBER: 60/119549
90 PRIOR FILING DATE: 1999-02-10
91 PRIOR APPLICATION NUMBER: 60/123618
92 PRIOR FILING DATE: 1999-03-10
93 PRIOR APPLICATION NUMBER: 60/125259
94 PRIOR FILING DATE: 1999-03-19
95 PRIOR APPLICATION NUMBER: 60/125775
96 PRIOR FILING DATE: 1999-03-23
97 PRIOR APPLICATION NUMBER: 60/126773
98 PRIOR FILING DATE: 1999-03-29
99 PRIOR APPLICATION NUMBER: 60/127887
100 PRIOR FILING DATE: 1999-04-05
101 PRIOR APPLICATION NUMBER: 60/130232
102 PRIOR FILING DATE: 1999-04-21
103 PRIOR APPLICATION NUMBER: 60/131022
104 PRIOR FILING DATE: 1999-04-26
105 PRIOR APPLICATION NUMBER: 60/131270
106 PRIOR FILING DATE: 1999-04-27
107 PRIOR APPLICATION NUMBER: 60/131291
108 PRIOR FILING DATE: 1999-04-27
109 PRIOR APPLICATION NUMBER: 60/131445
110 PRIOR FILING DATE: 1999-04-28
111 PRIOR APPLICATION NUMBER: 60/134287
112 PRIOR FILING DATE: 1999-05-14
113 PRIOR APPLICATION NUMBER: 60/140650
114 PRIOR FILING DATE: 1999-06-22
115 PRIOR APPLICATION NUMBER: 60/140723
116 PRIOR FILING DATE: 1999-06-22
117 PRIOR APPLICATION NUMBER: 60/141037
118 PRIOR FILING DATE: 1999-06-23
119 PRIOR APPLICATION NUMBER: 60/144758
120 PRIOR FILING DATE: 1999-07-20
121 PRIOR APPLICATION NUMBER: 60/145698
122 PRIOR FILING DATE: 1999-07-26
123 PRIOR APPLICATION NUMBER: 60/146222
124 PRIOR FILING DATE: 1999-07-28
125 PRIOR APPLICATION NUMBER: 60/146963
126 PRIOR FILING DATE: 1999-08-03
127 PRIOR APPLICATION NUMBER: 60/149320
128 PRIOR FILING DATE: 1999-08-17
129 PRIOR APPLICATION NUMBER: 60/149638
130 PRIOR FILING DATE: 1999-08-17
131 PRIOR APPLICATION NUMBER: 60/151733
132 PRIOR FILING DATE: 1999-08-31
133 PRIOR APPLICATION NUMBER: 60/164418
134 PRIOR FILING DATE: 1999-11-09
135 PRIOR APPLICATION NUMBER: 60/166361
136 PRIOR FILING DATE: 1999-11-16
137 PRIOR APPLICATION NUMBER: 60/169445
138 PRIOR FILING DATE: 1999-12-07
139 PRIOR APPLICATION NUMBER: 60/169495
140 PRIOR FILING DATE: 1999-12-07
141 PRIOR APPLICATION NUMBER: 60/169835

Query Match 1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3715 AATGAA 3762
Db 1729 AATGAA 1776

RESULT 48

US-10-219-076-39
; Sequence 39, Application US/10219076
; Publication No. US20030078379A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3530P1C62
; CURRENT APPLICATION NUMBER: US/10/219,076
; CURRENT FILING DATE: 2002-08-14
; PRIOR APPLICATION NUMBER: 10/119,480
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/081819
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081955
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082804
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/086392
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089905
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090691
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/095302
; PRIOR FILING DATE: 1998-08-04
; PRIOR APPLICATION NUMBER: 60/095318
; PRIOR FILING DATE: 1998-08-04
; PRIOR APPLICATION NUMBER: 60/095916
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: 60/096146
; PRIOR FILING DATE: 1998-08-11
; PRIOR APPLICATION NUMBER: 60/096791
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 60/097986
; PRIOR FILING DATE: 1998-08-26

US-10-230-434-39

US-10-230-434-39
; Sequence 39, Application US/10230434
; Publication No. US20030078380A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3530P1C62
; CURRENT APPLICATION NUMBER: US/10/219,076
; CURRENT FILING DATE: 2002-08-14
; PRIOR APPLICATION NUMBER: 10/119,480
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 39
; LENGTH: 1837
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-219-076-39

Query Match 1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3715 AATGAA 3762
Db 1729 AATGAA 1776

RESULT 49

US-10-230-434-39
; Sequence 39, Application US/10230434
; Publication No. US20030078380A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.

;; PRIOR APPLICATION NUMBER: 60/098544
;; PRIOR FILING DATE: 1998-08-31
;; PRIOR APPLICATION NUMBER: 60/099596
;; PRIOR FILING DATE: 1998-09-09
;; PRIOR APPLICATION NUMBER: 60/099598
;; PRIOR FILING DATE: 1998-09-09
;; PRIOR APPLICATION NUMBER: 60/099803
;; PRIOR FILING DATE: 1998-09-10
;; PRIOR APPLICATION NUMBER: 60/099811
;; PRIOR FILING DATE: 1998-09-10
;; PRIOR APPLICATION NUMBER: 60/099812
;; PRIOR FILING DATE: 1998-09-10
;; PRIOR APPLICATION NUMBER: 60/099816
;; PRIOR FILING DATE: 1998-09-10
;; PRIOR APPLICATION NUMBER: 60/100038
;; PRIOR FILING DATE: 1998-09-11
;; PRIOR APPLICATION NUMBER: 60/100385
;; PRIOR FILING DATE: 1998-09-15
;; PRIOR APPLICATION NUMBER: 60/100390
;; PRIOR FILING DATE: 1998-09-15
;; PRIOR APPLICATION NUMBER: 60/100627
;; PRIOR FILING DATE: 1998-09-16
;; PRIOR APPLICATION NUMBER: 60/100848
;; PRIOR FILING DATE: 1998-09-18
;; PRIOR APPLICATION NUMBER: 60/100919
;; PRIOR FILING DATE: 1998-09-17
;; PRIOR APPLICATION NUMBER: 60/101477
;; PRIOR FILING DATE: 1998-09-23
;; PRIOR APPLICATION NUMBER: 60/101738
;; PRIOR FILING DATE: 1998-09-24
;; PRIOR APPLICATION NUMBER: 60/101741
;; PRIOR FILING DATE: 1998-09-24
;; PRIOR APPLICATION NUMBER: 60/101786
;; PRIOR FILING DATE: 1998-09-25
;; PRIOR APPLICATION NUMBER: 60/101916
;; PRIOR FILING DATE: 1998-09-24
;; PRIOR APPLICATION NUMBER: 60/101922
;; PRIOR FILING DATE: 1998-09-24
;; PRIOR APPLICATION NUMBER: 60/106178
;; PRIOR FILING DATE: 1998-10-28
;; PRIOR APPLICATION NUMBER: 60/106248
;; PRIOR FILING DATE: 1998-10-29
;; PRIOR APPLICATION NUMBER: 60/106464
;; PRIOR FILING DATE: 1998-10-30
;; PRIOR APPLICATION NUMBER: 60/106905
;; PRIOR FILING DATE: 1998-11-03
;; PRIOR APPLICATION NUMBER: 60/108787
;; PRIOR FILING DATE: 1998-11-17
;; PRIOR APPLICATION NUMBER: 60/108901
;; PRIOR FILING DATE: 1998-11-17
;; PRIOR APPLICATION NUMBER: 60/108949
;; PRIOR FILING DATE: 1998-11-18
;; PRIOR APPLICATION NUMBER: 60/112422
;; PRIOR FILING DATE: 1998-12-15
;; PRIOR APPLICATION NUMBER: 60/113296
;; PRIOR FILING DATE: 1998-12-22
;; PRIOR APPLICATION NUMBER: 60/113605
;; PRIOR FILING DATE: 1998-12-23
;; PRIOR APPLICATION NUMBER: 60/113621
;; PRIOR FILING DATE: 1998-12-23
;; PRIOR APPLICATION NUMBER: 60/115558
;; PRIOR FILING DATE: 1999-01-12
;; PRIOR APPLICATION NUMBER: 60/115565
;; PRIOR FILING DATE: 1999-01-12
;; PRIOR APPLICATION NUMBER: 60/115733
;; PRIOR FILING DATE: 1999-01-12
;; PRIOR APPLICATION NUMBER: 60/119549
;; PRIOR FILING DATE: 1999-02-10
;; PRIOR APPLICATION NUMBER: 60/123618
;; PRIOR FILING DATE: 1999-03-10
;; PRIOR APPLICATION NUMBER: 60/125259
;; PRIOR FILING DATE: 1999-03-19
;; PRIOR APPLICATION NUMBER: 60/125775

;; PRIOR FILING DATE: 1999-03-23
;; PRIOR APPLICATION NUMBER: 60/126773
;; PRIOR FILING DATE: 1999-03-29
;; PRIOR APPLICATION NUMBER: 60/127887
;; PRIOR FILING DATE: 1999-04-05
;; PRIOR APPLICATION NUMBER: 60/130232
;; PRIOR FILING DATE: 1999-04-21
;; PRIOR APPLICATION NUMBER: 60/131022
;; PRIOR FILING DATE: 1999-04-26
;; PRIOR APPLICATION NUMBER: 60/131270
;; PRIOR FILING DATE: 1999-04-27
;; PRIOR APPLICATION NUMBER: 60/131291
;; PRIOR FILING DATE: 1999-04-27
;; PRIOR APPLICATION NUMBER: 60/131445
;; PRIOR FILING DATE: 1999-04-28
;; PRIOR APPLICATION NUMBER: 60/134287
;; PRIOR FILING DATE: 1999-05-14
;; PRIOR APPLICATION NUMBER: 60/140650
;; PRIOR FILING DATE: 1999-06-22
;; PRIOR APPLICATION NUMBER: 60/140723
;; PRIOR FILING DATE: 1999-06-22
;; PRIOR APPLICATION NUMBER: 60/141037
;; PRIOR FILING DATE: 1999-06-23
;; PRIOR APPLICATION NUMBER: 60/144758
;; PRIOR FILING DATE: 1999-07-20
;; PRIOR APPLICATION NUMBER: 60/145698
;; PRIOR FILING DATE: 1999-07-26
;; PRIOR APPLICATION NUMBER: 60/146222
;; PRIOR FILING DATE: 1999-07-28
;; PRIOR APPLICATION NUMBER: 60/146963
;; PRIOR FILING DATE: 1999-08-03
;; PRIOR APPLICATION NUMBER: 60/149320
;; PRIOR FILING DATE: 1999-08-17
;; PRIOR APPLICATION NUMBER: 60/149638
;; PRIOR FILING DATE: 1999-08-17
;; PRIOR APPLICATION NUMBER: 60/151733
;; PRIOR FILING DATE: 1999-08-31
;; PRIOR APPLICATION NUMBER: 60/164418
;; PRIOR FILING DATE: 1999-11-09
;; PRIOR APPLICATION NUMBER: 60/166361
;; PRIOR FILING DATE: 1999-11-16
;; PRIOR APPLICATION NUMBER: 60/169445
;; PRIOR FILING DATE: 1999-12-07
;; PRIOR APPLICATION NUMBER: 60/169495
;; PRIOR FILING DATE: 1999-12-07
;; PRIOR APPLICATION NUMBER: 60/169835

Query Match 1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Fred. No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3715 AATGAA 3762
Db 1729 AATGAA 1776

RESULT 50
US-10-219-003-39
; Sequence 39, Application US/10219003
; Publication No. US2003008063A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

TITLE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P3530P1C12
CURRENT APPLICATION NUMBER: US/10/219,003
CURRENT FILING DATE: 2002-08-12
PRIOR APPLICATION NUMBER: 10/119,480
PRIOR FILING DATE: 2002-04-09
PRIOR APPLICATION NUMBER: 60/059113
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/062287
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/063549
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/064103
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/069873
PRIOR FILING DATE: 1997-12-17
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/084441
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/085323
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/086392
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089905
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090691
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090695
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/095302
PRIOR FILING DATE: 1998-08-04
PRIOR APPLICATION NUMBER: 60/095318
PRIOR FILING DATE: 1998-08-04
PRIOR APPLICATION NUMBER: 60/095916
PRIOR FILING DATE: 1998-08-10
PRIOR APPLICATION NUMBER: 60/096146
PRIOR FILING DATE: 1998-08-11
PRIOR APPLICATION NUMBER: 60/096791
PRIOR FILING DATE: 1998-08-17
PRIOR APPLICATION NUMBER: 60/097986
PRIOR FILING DATE: 1998-08-26
PRIOR APPLICATION NUMBER: 60/098544
PRIOR FILING DATE: 1998-08-31
PRIOR APPLICATION NUMBER: 60/099596
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: 60/099598
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: 60/099803

PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099811
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099812
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099816
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/100038
PRIOR FILING DATE: 1998-09-11
PRIOR APPLICATION NUMBER: 60/100385
PRIOR FILING DATE: 1998-09-15
PRIOR APPLICATION NUMBER: 60/100390
PRIOR FILING DATE: 1998-09-15
PRIOR APPLICATION NUMBER: 60/100627
PRIOR FILING DATE: 1998-09-16
PRIOR APPLICATION NUMBER: 60/100848
PRIOR FILING DATE: 1998-09-18
PRIOR APPLICATION NUMBER: 60/100919
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: 60/101477
PRIOR FILING DATE: 1998-09-23
PRIOR APPLICATION NUMBER: 60/101738
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/101741
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/101786
PRIOR FILING DATE: 1998-09-25
PRIOR APPLICATION NUMBER: 60/101916
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/101922
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/106178
PRIOR FILING DATE: 1998-10-28
PRIOR APPLICATION NUMBER: 60/106248
PRIOR FILING DATE: 1998-10-29
PRIOR APPLICATION NUMBER: 60/106464
PRIOR FILING DATE: 1998-10-30
PRIOR APPLICATION NUMBER: 60/106905
PRIOR FILING DATE: 1998-11-03
PRIOR APPLICATION NUMBER: 60/108787
PRIOR FILING DATE: 1998-11-17
PRIOR APPLICATION NUMBER: 60/108801
PRIOR FILING DATE: 1998-11-17
PRIOR APPLICATION NUMBER: 60/108849
PRIOR FILING DATE: 1998-11-18
PRIOR APPLICATION NUMBER: 60/112422
PRIOR FILING DATE: 1998-12-15
PRIOR APPLICATION NUMBER: 60/113296
PRIOR FILING DATE: 1998-12-22
PRIOR APPLICATION NUMBER: 60/113605
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/113621
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/115558
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/115565
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/115733
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/119549
PRIOR FILING DATE: 1999-02-10
PRIOR APPLICATION NUMBER: 60/123618
PRIOR FILING DATE: 1999-03-10
PRIOR APPLICATION NUMBER: 60/125259
PRIOR FILING DATE: 1999-03-19
PRIOR APPLICATION NUMBER: 60/125775
PRIOR FILING DATE: 1999-03-23
PRIOR APPLICATION NUMBER: 60/126773
PRIOR FILING DATE: 1999-03-29
PRIOR APPLICATION NUMBER: 60/127887
PRIOR FILING DATE: 1999-04-05
PRIOR APPLICATION NUMBER: 60/130232
PRIOR FILING DATE: 1999-04-21

RESULT 51
US-10-219-075-39
Sequence 39, Application US/10219075
Publication No. US20030088064A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Desnoyers, Luc
APPLICANT: Gerritsen, Mary
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Smith, Victoria
APPLICANT: Stephan, Jean-Philippe F.
APPLICANT: Matanabe, Colin L.
APPLICANT: Wood, William I.
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC ACIDS ENCODING THE SAME
FILE REFERENCE: P3530P1G61
CURRENT APPLICATION NUMBER: US/10/219,075
CURRENT FILING DATE: 2002-08-14
PRIOR APPLICATION NUMBER: 10/119,480
PRIOR FILING DATE: 2002-04-09
PRIOR APPLICATION NUMBER: 60/059113

```

; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 39
; LENGTH: 1837
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-219-075-39

Query Match 1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Fred. No. 3.2e-11; Gaps 0;
Matches 48; Conservative 0; Mismatches 0; Indels 0;

QY 3715 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
    |||
Db 1729 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1776

RESULT 52
US-10-219-464-39
; Sequence 39, Application US/10219464
; Publication No. US20030088065A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEOTIDE SEQUENCES ENCODING THE SAME
; FILE REFERENCE: P35301C57
; CURRENT APPLICATION NUMBER: US/10/219,464
; CURRENT FILING DATE: 2002-08-14
; PRIOR APPLICATION NUMBER: 10/119,480
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26

```



```

; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 246

```

```
Query Match      1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

RESULT 53
US-10-219-466-39
; Sequence 39, Application US/10219466
; Publication NO. US20030088066A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME

Query Match 1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

D^b

1729 AATGTAAGTAAATAAATAAAAAAAA
|||||

D^b

1776

```

RESULT 54
US-10-219-479-39
; Sequence 39, Application US/10219479
; Publication No. US2003008067A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3530PIC33
; CURRENT APPLICATION NUMBER: US/10/219,479
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: 10/119,480
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; Remaining Prior application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 39
; LENGTH: 1837
; TYPE: DNA
; ORGANISM: Homo Sapien
; US-10-219-479-39

```

```
Query Match      1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.2e-11;
Matches 48: Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

RESULT 55
US-10-219-481-39
; Sequence 39, Application US/10219481
; Publication No. US2003008068A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Deenoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher

```
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3530PIC28
; CURRENT APPLICATION NUMBER: US/10/219,481
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: 10/119,480
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 39
; LENGTH: 1837
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-219-481-39

Query Match      1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3715 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 1729 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1776

RESULT 56
US-10-230-260-39
; Sequence 39, Application US/10230260
; Publication No. US20030088070A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3530PIC83
; CURRENT APPLICATION NUMBER: US/10/230,260
; CURRENT FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: 10/119,480
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 39
; LENGTH: 1837
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-219-481-39

Query Match      1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3715 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 1729 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1776

RESULT 56
US-10-230-260-39
; Sequence 39, Application US/10230260
; Publication No. US20030088070A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3530PIC83
; CURRENT APPLICATION NUMBER: US/10/230,260
; CURRENT FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: 10/119,480
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 39
; LENGTH: 1837
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-230-260-39

Query Match      1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3715 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 1729 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1776

RESULT 57
US-10-232-231-39
; Sequence 39, Application US/10232231
; Publication No. US20030088071A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3530PIC104
; CURRENT APPLICATION NUMBER: US/10/232,231
; CURRENT FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 10/119,480
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 39
; LENGTH: 1837
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-230-260-39
```

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 39
; LENGTH: 1837
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-232-231-39

Query Match 1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAA 3762
Db 1729 AATGAA 1776

RESULT 58

US-10-232-233-39
; Sequence 39, Application US/10232233
; Publication No. US2003008072A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3530PIC108

; CURRENT APPLICATION NUMBER: US/10/232,233

; CURRENT FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 10/119,480
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 39
; LENGTH: 1837
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-232-233-39

Query Match 1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAA 3762
Db 1729 AATGAA 1776

RESULT 59

US-10-216-165-39
; Sequence 39, Application US/10216165
; Publication No. US20030092886A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3530PIC7

; CURRENT APPLICATION NUMBER: US/10/216,165

; CURRENT FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/119,480
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 39
; LENGTH: 1837
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-216-165-39

Query Match 1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAA 3762
Db 1729 AATGAA 1776

RESULT 60

US-10-218-956-39
; Sequence 39, Application US/10218956
; Publication No. US20030092887A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria

APPLICANT: Stephan, Jean-Philippe F.
APPLICANT: Watanabe, Colin L.
APPLICANT: Wood, William I.
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3530P1C15
CURRENT FILING DATE: 2002-08-12
CURRENT FILING DATE: 2002-08-12
PRIOR APPLICATION NUMBER: 10/119,480
PRIOR FILING DATE: 2002-04-09
PRIOR APPLICATION NUMBER: 60/059113
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/062287
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/063549
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/064103
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/069873
PRIOR FILING DATE: 1997-12-17
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/084441
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/085323
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/086392
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089905
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090691
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090695
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/093302
PRIOR FILING DATE: 1998-08-04
PRIOR APPLICATION NUMBER: 60/093318
PRIOR FILING DATE: 1998-08-04
PRIOR APPLICATION NUMBER: 60/095916
PRIOR FILING DATE: 1998-08-10
PRIOR APPLICATION NUMBER: 60/096146
PRIOR FILING DATE: 1998-08-11
PRIOR APPLICATION NUMBER: 60/096791
PRIOR FILING DATE: 1998-08-17
PRIOR APPLICATION NUMBER: 60/097986
PRIOR FILING DATE: 1998-08-26
PRIOR APPLICATION NUMBER: 60/098544
PRIOR FILING DATE: 1998-08-31
PRIOR APPLICATION NUMBER: 60/099596
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: 60/099598
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: 60/099803
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099811
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099812
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099816
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/100038
PRIOR FILING DATE: 1998-09-11
PRIOR APPLICATION NUMBER: 60/100385
PRIOR FILING DATE: 1998-09-15
PRIOR APPLICATION NUMBER: 60/100390
PRIOR FILING DATE: 1998-09-15
PRIOR APPLICATION NUMBER: 60/100627
PRIOR FILING DATE: 1998-09-16
PRIOR APPLICATION NUMBER: 60/100848
PRIOR FILING DATE: 1998-09-18
PRIOR APPLICATION NUMBER: 60/100919
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: 60/101477
PRIOR FILING DATE: 1998-09-23
PRIOR APPLICATION NUMBER: 60/101738
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/101741
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/101786
PRIOR FILING DATE: 1998-09-25
PRIOR APPLICATION NUMBER: 60/101916
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/101922
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/106178
PRIOR FILING DATE: 1998-10-28
PRIOR APPLICATION NUMBER: 60/106248
PRIOR FILING DATE: 1998-10-29
PRIOR APPLICATION NUMBER: 60/106464
PRIOR FILING DATE: 1998-10-30
PRIOR APPLICATION NUMBER: 60/106905
PRIOR FILING DATE: 1998-11-03
PRIOR APPLICATION NUMBER: 60/108787
PRIOR FILING DATE: 1998-11-17
PRIOR APPLICATION NUMBER: 60/108801
PRIOR FILING DATE: 1998-11-17
PRIOR APPLICATION NUMBER: 60/108849
PRIOR FILING DATE: 1998-11-18
PRIOR APPLICATION NUMBER: 60/112422
PRIOR FILING DATE: 1998-12-15
PRIOR APPLICATION NUMBER: 60/113296
PRIOR FILING DATE: 1998-12-22
PRIOR APPLICATION NUMBER: 60/113605
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/113621
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/115558
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/115565
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/115733
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/119549
PRIOR FILING DATE: 1999-02-10
PRIOR APPLICATION NUMBER: 60/123618
PRIOR FILING DATE: 1999-03-10
PRIOR APPLICATION NUMBER: 60/125259
PRIOR FILING DATE: 1999-03-19
PRIOR APPLICATION NUMBER: 60/125775
PRIOR FILING DATE: 1999-03-23
PRIOR APPLICATION NUMBER: 60/126773
PRIOR FILING DATE: 1999-03-29

;; PRIOR APPLICATION NUMBER: 60/127887
;; PRIOR FILING DATE: 1999-04-05
;; PRIOR APPLICATION NUMBER: 60/130232
;; PRIOR FILING DATE: 1999-04-21
;; PRIOR APPLICATION NUMBER: 60/131022
;; PRIOR FILING DATE: 1999-04-26
;; PRIOR APPLICATION NUMBER: 60/131270
;; PRIOR FILING DATE: 1999-04-27
;; PRIOR APPLICATION NUMBER: 60/131291
;; PRIOR FILING DATE: 1999-04-27
;; PRIOR APPLICATION NUMBER: 60/131445
;; PRIOR FILING DATE: 1999-04-28
;; PRIOR APPLICATION NUMBER: 60/134287
;; PRIOR FILING DATE: 1999-05-14
;; PRIOR APPLICATION NUMBER: 60/140650
;; PRIOR FILING DATE: 1999-06-22
;; PRIOR APPLICATION NUMBER: 60/140723
;; PRIOR FILING DATE: 1999-06-22
;; PRIOR APPLICATION NUMBER: 60/141037
;; PRIOR FILING DATE: 1999-06-23
;; PRIOR APPLICATION NUMBER: 60/144758
;; PRIOR FILING DATE: 1999-07-20
;; PRIOR APPLICATION NUMBER: 60/145698
;; PRIOR FILING DATE: 1999-07-26
;; PRIOR APPLICATION NUMBER: 60/146222
;; PRIOR FILING DATE: 1999-07-28
;; PRIOR APPLICATION NUMBER: 60/146963
;; PRIOR FILING DATE: 1999-08-03
;; PRIOR APPLICATION NUMBER: 60/149320
;; PRIOR FILING DATE: 1999-08-17
;; PRIOR APPLICATION NUMBER: 60/149638
;; PRIOR FILING DATE: 1999-08-17
;; PRIOR APPLICATION NUMBER: 60/151733
;; PRIOR FILING DATE: 1999-08-31
;; PRIOR APPLICATION NUMBER: 60/164418
;; PRIOR FILING DATE: 1999-11-09
;; PRIOR APPLICATION NUMBER: 60/166361
;; PRIOR FILING DATE: 1999-11-16
;; PRIOR APPLICATION NUMBER: 60/169445
;; PRIOR FILING DATE: 1999-12-07
;; PRIOR APPLICATION NUMBER: 60/169495
;; PRIOR FILING DATE: 1999-12-07
;; PRIOR APPLICATION NUMBER: 60/169835

Query Match 1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3715 AATGAA 3762
Db 1729 AATGAA 1776

RESULT 61
US-10-219-468-39
; Sequence 39, Application US/10219468
; Publication No. US20030092888A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3530PIC34
; CURRENT APPLICATION NUMBER: US/10/219,468

;; CURRENT FILING DATE: 2002-08-13
;; PRIOR APPLICATION NUMBER: 10/119,480
;; PRIOR FILING DATE: 2002-04-09
;; PRIOR APPLICATION NUMBER: 60/059113
;; PRIOR FILING DATE: 1997-09-17
;; PRIOR APPLICATION NUMBER: 60/062287
;; PRIOR FILING DATE: 1997-10-17
;; PRIOR APPLICATION NUMBER: 60/063549
;; PRIOR FILING DATE: 1997-10-28
;; PRIOR APPLICATION NUMBER: 60/064103
;; PRIOR FILING DATE: 1997-10-31
;; PRIOR APPLICATION NUMBER: 60/069873
;; PRIOR FILING DATE: 1997-12-17
;; PRIOR APPLICATION NUMBER: 60/078910
;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/079294
;; PRIOR FILING DATE: 1998-03-25
;; PRIOR APPLICATION NUMBER: 60/079656
;; PRIOR FILING DATE: 1998-03-26
;; PRIOR APPLICATION NUMBER: 60/079728
;; PRIOR FILING DATE: 1998-03-27
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 246
;; SEQ ID NO 39
;; LENGTH: 1837
;; TYPE: DNA
;; ORGANISM: Homo Sapien
US-10-219-468-39

Query Match 1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3715 AATGAA 3762
Db 1729 AATGAA 1776

RESULT 62
US-10-219-478-39
; Sequence 39, Application US/10219478
; Publication No. US20030092889A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3530PIC30
; CURRENT APPLICATION NUMBER: US/10/219,478
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: 10/119,480
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20

```
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 39
; LENGTH: 1837
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-219-478-39

Query Match      1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
    |||||||
Db 1729 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1776

RESULT 63
US-10-219-536-39
; Sequence 39, Application US/10219536
; Publication No. US20030092890A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3530P1C67
; CURRENT APPLICATION NUMBER: US/10/219,536
; CURRENT FILING DATE: 2002-08-14
; PRIOR APPLICATION NUMBER: 10/119,480
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-03-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 39
; LENGTH: 1837
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-219-536-39

Query Match      1.3%; Score 48; DB 15; Length 1837;
```

```
Best Local Similarity 100.0%; Pred. No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
    |||||||
Db 1729 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1776

RESULT 64
US-10-233-205-39
; Sequence 39, Application US/10233205
; Publication No. US20030096362A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3530P1C115
; CURRENT APPLICATION NUMBER: US/10/233,205
; CURRENT FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 10/119,480
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-03-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 39
; LENGTH: 1837
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-233-205-39

Query Match      1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
    |||||||
Db 1729 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1776

RESULT 65
US-10-219-072-39
; Sequence 39, Application US/10219072
; Publication No. US20030096959A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
```

```

; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 39
; LENGTH: 1837
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-219-470-39

Query Match          1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.2e-11; Indels 0; Gaps 0;
Matches 48; Conservative 0; Mismatches 0;

QY 3715 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
    |||
Db 1729 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1776

RESULT 67
US-10-219-474-39
; Sequence 39, Application US/10219474
; Publication No. US20030096961A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3530PIC36
; CURRENT APPLICATION NUMBER: US/10/219,474
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: 10/119,480
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25

```


;; PRIOR APPLICATION NUMBER: 60/079656
;; PRIOR FILING DATE: 1998-03-26
;; PRIOR APPLICATION NUMBER: 60/079728
;; PRIOR FILING DATE: 1998-03-27
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 246
;; SEQ ID NO 39
;; LENGTH: 1837
;; TYPE: DNA
;; ORGANISM: Homo Sapien
US-10-219-524-39

Query Match 1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3715 AATGAAAAA 3762
Db 1729 AATGAAAAA 1776

RESULT 68

US-10-219-524-39
;; Sequence 39, Application US/10219524
;; Publication No. US20030096962A1
;; GENERAL INFORMATION:
;; APPLICANT: Baker, Kevin P.
;; APPLICANT: Desnoyers, Luc
;; APPLICANT: Gerritsen, Mary
;; APPLICANT: Goddard, Audrey
;; APPLICANT: Godowski, Paul J.
;; APPLICANT: Grimaldi, J. Christopher
;; APPLICANT: Gurney, Austin L.
;; APPLICANT: Smith, Victoria
;; APPLICANT: Stephan, Jean-Philippe F.
;; APPLICANT: Watanabe, Colin L.
;; APPLICANT: Wood, William I.
;; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;; FILE REFERENCE: P3530P1C37
;; CURRENT APPLICATION NUMBER: US/10/219,524
;; CURRENT FILING DATE: 2002-08-13
;; PRIOR APPLICATION NUMBER: 10/119,480
;; PRIOR FILING DATE: 2002-04-09
;; PRIOR APPLICATION NUMBER: 60/059113
;; PRIOR FILING DATE: 1997-09-17
;; PRIOR APPLICATION NUMBER: 60/062287
;; PRIOR FILING DATE: 1997-10-17
;; PRIOR APPLICATION NUMBER: 60/063549
;; PRIOR FILING DATE: 1997-10-28
;; PRIOR APPLICATION NUMBER: 60/064103
;; PRIOR FILING DATE: 1997-10-31
;; PRIOR APPLICATION NUMBER: 60/069873
;; PRIOR FILING DATE: 1997-12-17
;; PRIOR APPLICATION NUMBER: 60/078910
;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/079294
;; PRIOR FILING DATE: 1998-03-25
;; PRIOR APPLICATION NUMBER: 60/079656
;; PRIOR FILING DATE: 1998-03-26
;; PRIOR APPLICATION NUMBER: 60/079728
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/081819
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081955
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/082804
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/084441
;; PRIOR FILING DATE: 1998-05-06
;; PRIOR APPLICATION NUMBER: 60/085323
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085579
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/086392
;; PRIOR FILING DATE: 1998-05-22
;; PRIOR APPLICATION NUMBER: 60/089532
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089538
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089905
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/090472
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090557
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090691
;; PRIOR FILING DATE: 1998-06-25

Query Match 1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3715 AATGAAAAA 3762
Db 1729 AATGAAAAA 1776

RESULT 69
US-10-219-528-39
;; Sequence 39, Application US/10219528
;; Publication No. US20030096963A1
;; GENERAL INFORMATION:
;; APPLICANT: Baker, Kevin P.
;; APPLICANT: Desnoyers, Luc
;; APPLICANT: Gerritsen, Mary
;; APPLICANT: Goddard, Audrey
;; APPLICANT: Godowski, Paul J.
;; APPLICANT: Grimaldi, J. Christopher
;; APPLICANT: Gurney, Austin L.
;; APPLICANT: Smith, Victoria
;; APPLICANT: Stephan, Jean-Philippe F.
;; APPLICANT: Watanabe, Colin L.
;; APPLICANT: Wood, William I.
;; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;; FILE REFERENCE: P3530P1C88
;; CURRENT APPLICATION NUMBER: US/10/219,528
;; CURRENT FILING DATE: 2002-08-14
;; PRIOR APPLICATION NUMBER: 10/119,480
;; PRIOR FILING DATE: 2002-04-09
;; PRIOR APPLICATION NUMBER: 60/059113
;; PRIOR FILING DATE: 1997-09-17
;; PRIOR APPLICATION NUMBER: 60/062287
;; PRIOR FILING DATE: 1997-10-17
;; PRIOR APPLICATION NUMBER: 60/063549
;; PRIOR FILING DATE: 1997-10-28
;; PRIOR APPLICATION NUMBER: 60/064103
;; PRIOR FILING DATE: 1997-10-31
;; PRIOR APPLICATION NUMBER: 60/069873
;; PRIOR FILING DATE: 1997-12-17
;; PRIOR APPLICATION NUMBER: 60/078910
;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/079294
;; PRIOR FILING DATE: 1998-03-25
;; PRIOR APPLICATION NUMBER: 60/079656
;; PRIOR FILING DATE: 1998-03-26
;; PRIOR APPLICATION NUMBER: 60/079728
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/081819
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081955
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/082804
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/084441
;; PRIOR FILING DATE: 1998-05-06
;; PRIOR APPLICATION NUMBER: 60/085323
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085579
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/086392
;; PRIOR FILING DATE: 1998-05-22
;; PRIOR APPLICATION NUMBER: 60/089532
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089538
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089905
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/090472
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090557
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090691
;; PRIOR FILING DATE: 1998-06-25

;; PRIOR APPLICATION NUMBER: 60/090695
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/091982
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/095302
;; PRIOR FILING DATE: 1998-08-04
;; PRIOR APPLICATION NUMBER: 60/095318
;; PRIOR FILING DATE: 1998-08-04
;; PRIOR APPLICATION NUMBER: 60/095916
;; PRIOR FILING DATE: 1998-08-10
;; PRIOR APPLICATION NUMBER: 60/096146
;; PRIOR FILING DATE: 1998-08-11
;; PRIOR APPLICATION NUMBER: 60/096791
;; PRIOR FILING DATE: 1998-08-17
;; PRIOR APPLICATION NUMBER: 60/097986
;; PRIOR FILING DATE: 1998-08-26
;; PRIOR APPLICATION NUMBER: 60/098544
;; PRIOR FILING DATE: 1998-08-31
;; PRIOR APPLICATION NUMBER: 60/099596
;; PRIOR FILING DATE: 1998-09-09
;; PRIOR APPLICATION NUMBER: 60/099598
;; PRIOR FILING DATE: 1998-09-09
;; PRIOR APPLICATION NUMBER: 60/099803
;; PRIOR FILING DATE: 1998-09-10
;; PRIOR APPLICATION NUMBER: 60/099811
;; PRIOR FILING DATE: 1998-09-10
;; PRIOR APPLICATION NUMBER: 60/099812
;; PRIOR FILING DATE: 1998-09-10
;; PRIOR APPLICATION NUMBER: 60/099816
;; PRIOR FILING DATE: 1998-09-10
;; PRIOR APPLICATION NUMBER: 60/100038
;; PRIOR FILING DATE: 1998-09-11
;; PRIOR APPLICATION NUMBER: 60/100385
;; PRIOR FILING DATE: 1998-09-15
;; PRIOR APPLICATION NUMBER: 60/100390
;; PRIOR FILING DATE: 1998-09-15
;; PRIOR APPLICATION NUMBER: 60/100627
;; PRIOR FILING DATE: 1998-09-16
;; PRIOR APPLICATION NUMBER: 60/100848
;; PRIOR FILING DATE: 1998-09-18
;; PRIOR APPLICATION NUMBER: 60/100919
;; PRIOR FILING DATE: 1998-09-17
;; PRIOR APPLICATION NUMBER: 60/101477
;; PRIOR FILING DATE: 1998-09-23
;; PRIOR APPLICATION NUMBER: 60/101738
;; PRIOR FILING DATE: 1998-09-24
;; PRIOR APPLICATION NUMBER: 60/101741
;; PRIOR FILING DATE: 1998-09-24
;; PRIOR APPLICATION NUMBER: 60/101786
;; PRIOR FILING DATE: 1998-09-25
;; PRIOR APPLICATION NUMBER: 60/101916
;; PRIOR FILING DATE: 1998-09-24
;; PRIOR APPLICATION NUMBER: 60/101922
;; PRIOR FILING DATE: 1998-09-24
;; PRIOR APPLICATION NUMBER: 60/106178
;; PRIOR FILING DATE: 1998-10-28
;; PRIOR APPLICATION NUMBER: 60/106248
;; PRIOR FILING DATE: 1998-10-29
;; PRIOR APPLICATION NUMBER: 60/106464
;; PRIOR FILING DATE: 1998-10-30
;; PRIOR APPLICATION NUMBER: 60/106905
;; PRIOR FILING DATE: 1998-11-03
;; PRIOR APPLICATION NUMBER: 60/108787
;; PRIOR FILING DATE: 1998-11-17
;; PRIOR APPLICATION NUMBER: 60/108801
;; PRIOR FILING DATE: 1998-11-17
;; PRIOR APPLICATION NUMBER: 60/108849
;; PRIOR FILING DATE: 1998-11-18
;; PRIOR APPLICATION NUMBER: 60/112422
;; PRIOR FILING DATE: 1998-12-15
;; PRIOR APPLICATION NUMBER: 60/113296
;; PRIOR FILING DATE: 1998-12-22
;; PRIOR APPLICATION NUMBER: 60/113605

;; PRIOR FILING DATE: 1998-12-23
;; PRIOR APPLICATION NUMBER: 60/113621
;; PRIOR FILING DATE: 1998-12-23
;; PRIOR APPLICATION NUMBER: 60/115558
;; PRIOR FILING DATE: 1999-01-12
;; PRIOR APPLICATION NUMBER: 60/115565
;; PRIOR FILING DATE: 1999-01-12
;; PRIOR APPLICATION NUMBER: 60/115733
;; PRIOR FILING DATE: 1999-01-12
;; PRIOR APPLICATION NUMBER: 60/119549
;; PRIOR FILING DATE: 1999-02-10
;; PRIOR APPLICATION NUMBER: 60/123618
;; PRIOR FILING DATE: 1999-03-10
;; PRIOR APPLICATION NUMBER: 60/125259
;; PRIOR FILING DATE: 1999-03-19
;; PRIOR APPLICATION NUMBER: 60/125775
;; PRIOR FILING DATE: 1999-03-23
;; PRIOR APPLICATION NUMBER: 60/126773
;; PRIOR FILING DATE: 1999-03-29
;; PRIOR APPLICATION NUMBER: 60/127887
;; PRIOR FILING DATE: 1999-04-05
;; PRIOR APPLICATION NUMBER: 60/130232
;; PRIOR FILING DATE: 1999-04-21
;; PRIOR APPLICATION NUMBER: 60/131022
;; PRIOR FILING DATE: 1999-04-26
;; PRIOR APPLICATION NUMBER: 60/131270
;; PRIOR FILING DATE: 1999-04-27
;; PRIOR APPLICATION NUMBER: 60/131291
;; PRIOR FILING DATE: 1999-04-27
;; PRIOR APPLICATION NUMBER: 60/131445
;; PRIOR FILING DATE: 1999-04-28
;; PRIOR APPLICATION NUMBER: 60/134287
;; PRIOR FILING DATE: 1999-05-14
;; PRIOR APPLICATION NUMBER: 60/140650
;; PRIOR FILING DATE: 1999-06-22
;; PRIOR APPLICATION NUMBER: 60/140723
;; PRIOR FILING DATE: 1999-06-22
;; PRIOR APPLICATION NUMBER: 60/141037
;; PRIOR FILING DATE: 1999-06-23
;; PRIOR APPLICATION NUMBER: 60/144758
;; PRIOR FILING DATE: 1999-07-20
;; PRIOR APPLICATION NUMBER: 60/145698
;; PRIOR FILING DATE: 1999-07-26
;; PRIOR APPLICATION NUMBER: 60/146222
;; PRIOR FILING DATE: 1999-07-28
;; PRIOR APPLICATION NUMBER: 60/146963
;; PRIOR FILING DATE: 1999-08-03
;; PRIOR APPLICATION NUMBER: 60/149320
;; PRIOR FILING DATE: 1999-08-17
;; PRIOR APPLICATION NUMBER: 60/149638
;; PRIOR FILING DATE: 1999-08-17
;; PRIOR APPLICATION NUMBER: 60/151733
;; PRIOR FILING DATE: 1999-08-31
;; PRIOR APPLICATION NUMBER: 60/164418
;; PRIOR FILING DATE: 1999-11-09
;; PRIOR APPLICATION NUMBER: 60/166361
;; PRIOR FILING DATE: 1999-11-16
;; PRIOR APPLICATION NUMBER: 60/169445
;; PRIOR FILING DATE: 1999-12-07
;; PRIOR APPLICATION NUMBER: 60/169495
;; PRIOR FILING DATE: 1999-12-07
;; PRIOR APPLICATION NUMBER: 60/169835

Query Match 1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3715 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 1729 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1776

RESULT 70

US-10-227-880-39
; Sequence 39, Application US/10227880
; Publication No. US20030096964A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gunney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: ACIDIC AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3530P1C74
; CURRENT APPLICATION NUMBER: US/10/227,880
; PRIOR FILING DATE: 2002-08-26
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/081819
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081955
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082804
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/086392
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089905
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090691
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/095302
; PRIOR FILING DATE: 1998-08-04
; PRIOR APPLICATION NUMBER: 60/095318
; PRIOR FILING DATE: 1998-08-04
; PRIOR APPLICATION NUMBER: 60/095916
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: 60/096146
; PRIOR FILING DATE: 1998-08-11
; PRIOR APPLICATION NUMBER: 60/096791
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 60/097986
; PRIOR FILING DATE: 1998-08-26
; PRIOR APPLICATION NUMBER: 60/098544
; PRIOR FILING DATE: 1998-08-31
; PRIOR APPLICATION NUMBER: 60/099596
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099598
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099803
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099811
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099812
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099816
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/100038
; PRIOR FILING DATE: 1998-09-11
; PRIOR APPLICATION NUMBER: 60/100385
; PRIOR FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: 60/100390
; PRIOR FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: 60/100627
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: 60/100848
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/100919
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: 60/101477
; PRIOR FILING DATE: 1998-09-23
; PRIOR APPLICATION NUMBER: 60/101738
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/101741
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/101786
; PRIOR FILING DATE: 1998-09-25
; PRIOR APPLICATION NUMBER: 60/101916
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/101922
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/106178
; PRIOR FILING DATE: 1998-10-28
; PRIOR APPLICATION NUMBER: 60/106248
; PRIOR FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: 60/106464
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: 60/106905
; PRIOR FILING DATE: 1998-11-03
; PRIOR APPLICATION NUMBER: 60/108787
; PRIOR FILING DATE: 1998-11-17
; PRIOR APPLICATION NUMBER: 60/108801
; PRIOR FILING DATE: 1998-11-17
; PRIOR APPLICATION NUMBER: 60/108849
; PRIOR FILING DATE: 1998-11-18
; PRIOR APPLICATION NUMBER: 60/112422
; PRIOR FILING DATE: 1998-12-15
; PRIOR APPLICATION NUMBER: 60/113296
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: 60/113605
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: 60/113621
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: 60/115558
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/115565
; PRIOR FILING DATE: 1999-01-12

PRIOR APPLICATION NUMBER: 60/115733
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/119549
PRIOR FILING DATE: 1999-02-10
PRIOR APPLICATION NUMBER: 60/123618
PRIOR FILING DATE: 1999-03-10
PRIOR APPLICATION NUMBER: 60/125259
PRIOR FILING DATE: 1999-03-19
PRIOR APPLICATION NUMBER: 60/125775
PRIOR FILING DATE: 1999-03-23
PRIOR APPLICATION NUMBER: 60/126773
PRIOR FILING DATE: 1999-03-29
PRIOR APPLICATION NUMBER: 60/127887
PRIOR FILING DATE: 1999-04-05
PRIOR APPLICATION NUMBER: 60/130232
PRIOR FILING DATE: 1999-04-21
PRIOR APPLICATION NUMBER: 60/131022
PRIOR FILING DATE: 1999-04-26
PRIOR APPLICATION NUMBER: 60/131270
PRIOR FILING DATE: 1999-04-27
PRIOR APPLICATION NUMBER: 60/131291
PRIOR FILING DATE: 1999-04-27
PRIOR APPLICATION NUMBER: 60/131445
PRIOR FILING DATE: 1999-04-28
PRIOR APPLICATION NUMBER: 60/134287
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: 60/140650
PRIOR FILING DATE: 1999-06-22
PRIOR APPLICATION NUMBER: 60/140723
PRIOR FILING DATE: 1999-06-22
PRIOR APPLICATION NUMBER: 60/141037
PRIOR FILING DATE: 1999-06-23
PRIOR APPLICATION NUMBER: 60/144758
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/145698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: 60/146222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: 60/146963
PRIOR FILING DATE: 1999-08-03
PRIOR APPLICATION NUMBER: 60/149320
PRIOR FILING DATE: 1999-08-17
PRIOR APPLICATION NUMBER: 60/149638
PRIOR FILING DATE: 1999-08-17
PRIOR APPLICATION NUMBER: 60/151733
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: 60/164418
PRIOR FILING DATE: 1999-11-09
PRIOR APPLICATION NUMBER: 60/166361
PRIOR FILING DATE: 1999-11-16
PRIOR APPLICATION NUMBER: 60/169445
PRIOR FILING DATE: 1999-12-07
PRIOR APPLICATION NUMBER: 60/169495
PRIOR FILING DATE: 1999-12-07
PRIOR APPLICATION NUMBER: 60/169835

Query Match 1.3%; Score 48; DB 15; Length 1837;

Best Local Similarity 100.0%; Pred.No. 3.2e-11;

Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAA 3762
DB 1729 AATGAA 1776

RESULT 71

US-10-227-881-39

Sequence 39, Application US/10227881

Publication No. US20030096965A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.

APPLICANT: Desnoyers, Luc

APPLICANT: Gerritsen, Mary

APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Smith, Victoria
APPLICANT: Stephan, Jean-Philippe F.
APPLICANT: Watanabe, Colin L.
APPLICANT: Wood, William I.
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3530PLC80
CURRENT APPLICATION NUMBER: US/10/227,881
CURRENT FILING DATE: 2002-08-26
PRIOR APPLICATION NUMBER: 10/119,480
PRIOR FILING DATE: 2002-04-09
PRIOR APPLICATION NUMBER: 60/059113
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/062287
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/063549
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/064103
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/069873
PRIOR FILING DATE: 1997-12-17
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/084441
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/085323
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/086392
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089905
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090691
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090695
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/095302
PRIOR FILING DATE: 1998-08-04
PRIOR APPLICATION NUMBER: 60/095318
PRIOR FILING DATE: 1998-08-04
PRIOR APPLICATION NUMBER: 60/095916
PRIOR FILING DATE: 1998-08-10
PRIOR APPLICATION NUMBER: 60/096146
PRIOR FILING DATE: 1998-08-11
PRIOR APPLICATION NUMBER: 60/096791
PRIOR FILING DATE: 1998-08-17

;; PRIOR APPLICATION NUMBER: 60/097986
;; PRIOR FILING DATE: 1998-08-26
;; PRIOR APPLICATION NUMBER: 60/098544
;; PRIOR FILING DATE: 1998-08-31
;; PRIOR APPLICATION NUMBER: 60/099596
;; PRIOR FILING DATE: 1998-09-09
;; PRIOR APPLICATION NUMBER: 60/099598
;; PRIOR FILING DATE: 1998-09-09
;; PRIOR APPLICATION NUMBER: 60/099803
;; PRIOR FILING DATE: 1998-09-10
;; PRIOR APPLICATION NUMBER: 60/099811
;; PRIOR FILING DATE: 1998-09-10
;; PRIOR APPLICATION NUMBER: 60/099812
;; PRIOR FILING DATE: 1998-09-10
;; PRIOR APPLICATION NUMBER: 60/099816
;; PRIOR FILING DATE: 1998-09-10
;; PRIOR APPLICATION NUMBER: 60/100038
;; PRIOR FILING DATE: 1998-09-11
;; PRIOR APPLICATION NUMBER: 60/100385
;; PRIOR FILING DATE: 1998-09-15
;; PRIOR APPLICATION NUMBER: 60/100390
;; PRIOR FILING DATE: 1998-09-15
;; PRIOR APPLICATION NUMBER: 60/100627
;; PRIOR FILING DATE: 1998-09-16
;; PRIOR APPLICATION NUMBER: 60/100848
;; PRIOR FILING DATE: 1998-09-18
;; PRIOR APPLICATION NUMBER: 60/100919
;; PRIOR FILING DATE: 1998-09-17
;; PRIOR APPLICATION NUMBER: 60/101477
;; PRIOR FILING DATE: 1998-09-23
;; PRIOR APPLICATION NUMBER: 60/101738
;; PRIOR FILING DATE: 1998-09-24
;; PRIOR APPLICATION NUMBER: 60/101741
;; PRIOR FILING DATE: 1998-09-24
;; PRIOR APPLICATION NUMBER: 60/101786
;; PRIOR FILING DATE: 1998-09-25
;; PRIOR APPLICATION NUMBER: 60/101916
;; PRIOR FILING DATE: 1998-09-24
;; PRIOR APPLICATION NUMBER: 60/101922
;; PRIOR FILING DATE: 1998-09-24
;; PRIOR APPLICATION NUMBER: 60/106178
;; PRIOR FILING DATE: 1998-10-28
;; PRIOR APPLICATION NUMBER: 60/106248
;; PRIOR FILING DATE: 1998-10-29
;; PRIOR APPLICATION NUMBER: 60/106464
;; PRIOR FILING DATE: 1998-10-30
;; PRIOR APPLICATION NUMBER: 60/106905
;; PRIOR FILING DATE: 1998-11-03
;; PRIOR APPLICATION NUMBER: 60/108787
;; PRIOR FILING DATE: 1998-11-17
;; PRIOR APPLICATION NUMBER: 60/108801
;; PRIOR FILING DATE: 1998-11-17
;; PRIOR APPLICATION NUMBER: 60/108849
;; PRIOR FILING DATE: 1998-11-18
;; PRIOR APPLICATION NUMBER: 60/112422
;; PRIOR FILING DATE: 1998-12-15
;; PRIOR APPLICATION NUMBER: 60/113296
;; PRIOR FILING DATE: 1998-12-22
;; PRIOR APPLICATION NUMBER: 60/113605
;; PRIOR FILING DATE: 1998-12-23
;; PRIOR APPLICATION NUMBER: 60/113621
;; PRIOR FILING DATE: 1998-12-23
;; PRIOR APPLICATION NUMBER: 60/115558
;; PRIOR FILING DATE: 1999-01-12
;; PRIOR APPLICATION NUMBER: 60/115565
;; PRIOR FILING DATE: 1999-01-12
;; PRIOR APPLICATION NUMBER: 60/115733
;; PRIOR FILING DATE: 1999-01-12
;; PRIOR APPLICATION NUMBER: 60/119549
;; PRIOR FILING DATE: 1999-02-10
;; PRIOR APPLICATION NUMBER: 60/123618
;; PRIOR FILING DATE: 1999-03-10
;; PRIOR APPLICATION NUMBER: 60/125259

;; PRIOR FILING DATE: 1999-03-19
;; PRIOR APPLICATION NUMBER: 60/125775
;; PRIOR FILING DATE: 1999-03-23
;; PRIOR APPLICATION NUMBER: 60/126773
;; PRIOR FILING DATE: 1999-03-29
;; PRIOR APPLICATION NUMBER: 60/127887
;; PRIOR FILING DATE: 1999-04-05
;; PRIOR APPLICATION NUMBER: 60/130232
;; PRIOR FILING DATE: 1999-04-21
;; PRIOR APPLICATION NUMBER: 60/131022
;; PRIOR FILING DATE: 1999-04-26
;; PRIOR APPLICATION NUMBER: 60/131270
;; PRIOR FILING DATE: 1999-04-27
;; PRIOR APPLICATION NUMBER: 60/131291
;; PRIOR FILING DATE: 1999-04-27
;; PRIOR APPLICATION NUMBER: 60/131445
;; PRIOR FILING DATE: 1999-04-28
;; PRIOR APPLICATION NUMBER: 60/134287
;; PRIOR FILING DATE: 1999-05-14
;; PRIOR APPLICATION NUMBER: 60/140650
;; PRIOR FILING DATE: 1999-06-22
;; PRIOR APPLICATION NUMBER: 60/140723
;; PRIOR FILING DATE: 1999-06-22
;; PRIOR APPLICATION NUMBER: 60/141037
;; PRIOR FILING DATE: 1999-06-23
;; PRIOR APPLICATION NUMBER: 60/144758
;; PRIOR FILING DATE: 1999-07-20
;; PRIOR APPLICATION NUMBER: 60/145698
;; PRIOR FILING DATE: 1999-07-26
;; PRIOR APPLICATION NUMBER: 60/146222
;; PRIOR FILING DATE: 1999-07-28
;; PRIOR APPLICATION NUMBER: 60/146963
;; PRIOR FILING DATE: 1999-08-03
;; PRIOR APPLICATION NUMBER: 60/149320
;; PRIOR FILING DATE: 1999-08-17
;; PRIOR APPLICATION NUMBER: 60/149638
;; PRIOR FILING DATE: 1999-08-17
;; PRIOR APPLICATION NUMBER: 60/151733
;; PRIOR FILING DATE: 1999-08-31
;; PRIOR APPLICATION NUMBER: 60/164418
;; PRIOR FILING DATE: 1999-11-09
;; PRIOR APPLICATION NUMBER: 60/166361
;; PRIOR FILING DATE: 1999-11-16
;; PRIOR APPLICATION NUMBER: 60/169445
;; PRIOR FILING DATE: 1999-12-07
;; PRIOR APPLICATION NUMBER: 60/169495
;; PRIOR FILING DATE: 1999-12-07
;; PRIOR APPLICATION NUMBER: 60/169835

Query Match 1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3715 AATGAA 3762
|||||
Db 1729 AATGAA 1776
|||||

RESULT 72
US-10-227-882-39
; Sequence 39, Application US/10227882
; Publication No. US20030096966A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.

APPLICANT: Wood, William I.
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3530P1C81
CURRENT APPLICATION NUMBER: US/10/227,882
PRIOR FILING DATE: 2002-08-26
PRIOR APPLICATION NUMBER: 10/119,480
PRIOR FILING DATE: 2002-04-09
PRIOR APPLICATION NUMBER: 60/039113
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/062287
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/063549
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/064103
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/069873
PRIOR FILING DATE: 1997-12-17
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/084441
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/085323
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/086392
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089905
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090691
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090695
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/095302
PRIOR FILING DATE: 1998-08-04
PRIOR APPLICATION NUMBER: 60/095318
PRIOR FILING DATE: 1998-08-04
PRIOR APPLICATION NUMBER: 60/095916
PRIOR FILING DATE: 1998-08-10
PRIOR APPLICATION NUMBER: 60/096146
PRIOR FILING DATE: 1998-08-11
PRIOR APPLICATION NUMBER: 60/096791
PRIOR FILING DATE: 1998-08-17
PRIOR APPLICATION NUMBER: 60/097986
PRIOR FILING DATE: 1998-08-26
PRIOR APPLICATION NUMBER: 60/098544
PRIOR FILING DATE: 1998-08-31
PRIOR APPLICATION NUMBER: 60/099596
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: 60/099598
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: 60/099803
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099811
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099812
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099816
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/100038
PRIOR FILING DATE: 1998-09-11
PRIOR APPLICATION NUMBER: 60/100385
PRIOR FILING DATE: 1998-09-15
PRIOR APPLICATION NUMBER: 60/100390
PRIOR FILING DATE: 1998-09-15
PRIOR APPLICATION NUMBER: 60/100627
PRIOR FILING DATE: 1998-09-16
PRIOR APPLICATION NUMBER: 60/100848
PRIOR FILING DATE: 1998-09-18
PRIOR APPLICATION NUMBER: 60/100919
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: 60/101477
PRIOR FILING DATE: 1998-09-23
PRIOR APPLICATION NUMBER: 60/101738
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/101741
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/101786
PRIOR FILING DATE: 1998-09-25
PRIOR APPLICATION NUMBER: 60/101916
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/101922
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/106178
PRIOR FILING DATE: 1998-10-28
PRIOR APPLICATION NUMBER: 60/106248
PRIOR FILING DATE: 1998-10-29
PRIOR APPLICATION NUMBER: 60/106464
PRIOR FILING DATE: 1998-10-30
PRIOR APPLICATION NUMBER: 60/106905
PRIOR FILING DATE: 1998-11-03
PRIOR APPLICATION NUMBER: 60/108787
PRIOR FILING DATE: 1998-11-17
PRIOR APPLICATION NUMBER: 60/108801
PRIOR FILING DATE: 1998-11-17
PRIOR APPLICATION NUMBER: 60/108849
PRIOR FILING DATE: 1998-11-18
PRIOR APPLICATION NUMBER: 60/112422
PRIOR FILING DATE: 1998-12-15
PRIOR APPLICATION NUMBER: 60/113296
PRIOR FILING DATE: 1998-12-22
PRIOR APPLICATION NUMBER: 60/113605
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/113621
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/115558
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/115565
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/115733
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/119549
PRIOR FILING DATE: 1999-02-10
PRIOR APPLICATION NUMBER: 60/123618
PRIOR FILING DATE: 1999-03-10
PRIOR APPLICATION NUMBER: 60/125259
PRIOR FILING DATE: 1999-03-19
PRIOR APPLICATION NUMBER: 60/125775
PRIOR FILING DATE: 1999-03-23
PRIOR APPLICATION NUMBER: 60/126773
PRIOR FILING DATE: 1999-03-29
PRIOR APPLICATION NUMBER: 60/127887
PRIOR FILING DATE: 1999-04-05

```
; PRIOR APPLICATION NUMBER: 60/130232
; PRIOR FILING DATE: 1999-04-21
; PRIOR APPLICATION NUMBER: 60/131022
; PRIOR FILING DATE: 1999-04-26
; PRIOR APPLICATION NUMBER: 60/131270
; PRIOR FILING DATE: 1999-04-27
; PRIOR APPLICATION NUMBER: 60/131291
; PRIOR FILING DATE: 1999-04-27
; PRIOR APPLICATION NUMBER: 60/131445
; PRIOR FILING DATE: 1999-04-28
; PRIOR APPLICATION NUMBER: 60/134287
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/140650
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: 60/140723
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: 60/141037
; PRIOR FILING DATE: 1999-06-23
; PRIOR APPLICATION NUMBER: 60/144758
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 60/145698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: 60/146222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: 60/146963
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: 60/149320
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: 60/149638
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: 60/151733
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: 60/164418
; PRIOR FILING DATE: 1999-11-09
; PRIOR APPLICATION NUMBER: 60/166361
; PRIOR FILING DATE: 1999-11-15
; PRIOR APPLICATION NUMBER: 60/169445
; PRIOR FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: 60/169495
; PRIOR FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: 60/169835
```

Query Match 1.3%; Score 48; DB 15; Length 1837;

Best Local Similarity 100.0%; Pred. No. 3.2e-11; Mismatches 0; Indels 0; Gaps 0;

```
QY 3715 ATGAAAAA 3762
Db 1729 AATGAAAA 1776
```

RESULT 73

```
US-10-230-436-39
; Sequence 39, Application US/10230436
; Publication No. US20030096761
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3530P1C97
; CURRENT FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: US/10/230,436
```

```
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 39
; LENGTH: 1837
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-230-436-39
```

Query Match 1.3%; Score 48; DB 15; Length 1837;

Best Local Similarity 100.0%; Pred. No. 3.2e-11; Mismatches 0; Indels 0; Gaps 0;

```
QY 3715 ATGAAAAA 3762
Db 1729 AATGAAAA 1776
```

RESULT 74

```
US-10-232-223-39
; Sequence 39, Application US/10232223
; Publication No. US20030096681
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3530P1C102
; CURRENT FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: US/10/232,223
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
```



```
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3530PIC105
; CURRENT APPLICATION NUMBER: US/10/232,229
; CURRENT FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 10/119,480
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3530PIC106
; CURRENT APPLICATION NUMBER: US/10/232,234
; CURRENT FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 10/119,480
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3530PIC106
; CURRENT APPLICATION NUMBER: US/10/232,234
; CURRENT FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 10/119,480
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
```

```
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 39
; LENGTH: 1837
; TYPE: DNA
; ORGANISM: Homo Sapien
; US-10-232-234-39

Query Match 1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
DB 1729 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1776

RESULT 79
US-10-219-060-39
; Sequence 39, Application US/10219060
; Publication No. US20030100064A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3530PIC22
; CURRENT APPLICATION NUMBER: US/10/219,060
; CURRENT FILING DATE: 2002-08-12
; PRIOR APPLICATION NUMBER: 10/119,480
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
```


;; PRIOR APPLICATION NUMBER: 60/146963
;; PRIOR FILING DATE: 1999-08-03
;; PRIOR APPLICATION NUMBER: 60/149320
;; PRIOR FILING DATE: 1999-08-17
;; PRIOR APPLICATION NUMBER: 60/149638
;; PRIOR FILING DATE: 1999-08-17
;; PRIOR APPLICATION NUMBER: 60/151733
;; PRIOR FILING DATE: 1999-08-31
;; PRIOR APPLICATION NUMBER: 60/164418
;; PRIOR FILING DATE: 1999-11-09
;; PRIOR APPLICATION NUMBER: 60/166361
;; PRIOR FILING DATE: 1999-11-16
;; PRIOR APPLICATION NUMBER: 60/169445
;; PRIOR FILING DATE: 1999-12-07
;; PRIOR APPLICATION NUMBER: 60/169495
;; PRIOR FILING DATE: 1999-12-07
;; PRIOR APPLICATION NUMBER: 60/169835

Query Match 1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAA 3762
|||||
DB 1729 AATGAA 1776

RESULT 80
US-10-216-160-39
; Sequence 39, Application US/10216160
; Publication No. US20030100708A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3530PIC8
; CURRENT APPLICATION NUMBER: US/10/216.160
; CURRENT FILING DATE: 2002-08-09
; Remaining Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 39
; LENGTH: 1837
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-216-160-39

Query Match 1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAA 3762
|||||
DB 1729 AATGAA 1776

RESULT 81
US-10-216-162-39
; Sequence 39, Application US/10216162
; Publication No. US20030100709A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary

;; APPLICANT: Goddard, Audrey
;; APPLICANT: Godowski, Paul J.
;; APPLICANT: Grimaldi, J. Christopher
;; APPLICANT: Gurney, Austin L.
;; APPLICANT: Smith, Victoria
;; APPLICANT: Stephan, Jean-Philippe F.
;; APPLICANT: Watanabe, Colin L.
;; APPLICANT: Wood, William I.
;; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;; FILE REFERENCE: P3530PIC2
;; CURRENT APPLICATION NUMBER: US/10/216.162
;; CURRENT FILING DATE: 2002-08-09
;; PRIOR APPLICATION NUMBER: 10/119.480
;; PRIOR FILING DATE: 2002-04-09
;; PRIOR APPLICATION NUMBER: 60/059113
;; PRIOR FILING DATE: 1997-09-17
;; PRIOR APPLICATION NUMBER: 60/062287
;; PRIOR FILING DATE: 1997-10-17
;; PRIOR APPLICATION NUMBER: 60/063549
;; PRIOR FILING DATE: 1997-10-28
;; PRIOR APPLICATION NUMBER: 60/064103
;; PRIOR FILING DATE: 1997-10-31
;; PRIOR APPLICATION NUMBER: 60/069873
;; PRIOR FILING DATE: 1997-12-17
;; PRIOR APPLICATION NUMBER: 60/078910
;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/079294
;; PRIOR FILING DATE: 1998-03-25
;; PRIOR APPLICATION NUMBER: 60/079656
;; PRIOR FILING DATE: 1998-03-26
;; PRIOR APPLICATION NUMBER: 60/079728
;; PRIOR FILING DATE: 1998-03-27
;; Remaining Prior Application data removed - See File Wrapper or PALM.
; SEQ ID NO 39
; NUMBER OF SEQ ID NOS: 246
; LENGTH: 1837
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-216-162-39

Query Match 1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAA 3762
|||||
DB 1729 AATGAA 1776

RESULT 82
US-10-216-164-39
; Sequence 39, Application US/10216164
; Publication No. US20030100710A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3530PIC5
; CURRENT APPLICATION NUMBER: US/10/216.164
; CURRENT FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/119.480
; PRIOR FILING DATE: 2002-04-09

;; PRIOR APPLICATION NUMBER: 60/059113
;; PRIOR FILING DATE: 1997-09-17
;; PRIOR APPLICATION NUMBER: 60/062287
;; PRIOR FILING DATE: 1997-10-17
;; PRIOR APPLICATION NUMBER: 60/063549
;; PRIOR FILING DATE: 1997-10-28
;; PRIOR APPLICATION NUMBER: 60/064103
;; PRIOR FILING DATE: 1997-10-31
;; PRIOR APPLICATION NUMBER: 60/069873
;; PRIOR FILING DATE: 1997-12-17
;; PRIOR APPLICATION NUMBER: 60/078910
;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/079294
;; PRIOR FILING DATE: 1998-03-25
;; PRIOR APPLICATION NUMBER: 60/079656
;; PRIOR FILING DATE: 1998-03-26
;; PRIOR APPLICATION NUMBER: 60/079728
;; PRIOR FILING DATE: 1998-03-27
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 246
;; SEQ ID NO 39
;; LENGTH: 1837
;; TYPE: DNA
;; ORGANISM: Homo Sapien
US-10-216-164-39

Query Match 1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAAAAA 1776
Db 1729 AATGAAAAA 1776

RESULT 83
US-10-216-167-39
;; Sequence 39, Application US/10216167
;; Publication No. US20030100711A1
;; GENERAL INFORMATION:
;; APPLICANT: Baker, Kevin P.
;; APPLICANT: Desnoyers, Luc
;; APPLICANT: Gerritsen, Mary
;; APPLICANT: Goddard, Audrey
;; APPLICANT: Godowski, Paul J.
;; APPLICANT: Grimaldi, J. Christopher
;; APPLICANT: Gurney, Austin L.
;; APPLICANT: Smith, Victoria
;; APPLICANT: Stephan, Jean-Philippe F.
;; APPLICANT: Watanabe, Colin L.
;; APPLICANT: Wood, William I.
;; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;; FILE REFERENCE: P3530PIC4
;; CURRENT FILING DATE: 2002-08-09
;; PRIOR APPLICATION NUMBER: 60/059113
;; PRIOR FILING DATE: 1997-09-17
;; PRIOR APPLICATION NUMBER: 60/062287
;; PRIOR FILING DATE: 1997-10-17
;; PRIOR APPLICATION NUMBER: 60/063549
;; PRIOR FILING DATE: 1997-10-28
;; PRIOR APPLICATION NUMBER: 60/064103
;; PRIOR FILING DATE: 1997-10-31
;; PRIOR APPLICATION NUMBER: 60/069873
;; PRIOR FILING DATE: 1997-12-17
;; PRIOR APPLICATION NUMBER: 60/078910
;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/079294
;; PRIOR FILING DATE: 1998-03-25
;; PRIOR APPLICATION NUMBER: 60/079656
;; PRIOR FILING DATE: 1998-03-26
;; PRIOR APPLICATION NUMBER: 60/079728
;; PRIOR FILING DATE: 1998-03-27
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 246
;; SEQ ID NO 39
;; LENGTH: 1837
;; TYPE: DNA
;; ORGANISM: Homo Sapien
US-10-216-164-39

;; PRIOR FILING DATE: 1998-03-26
;; PRIOR APPLICATION NUMBER: 60/079728
;; PRIOR FILING DATE: 1998-03-27
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 246
;; SEQ ID NO 39
;; LENGTH: 1837
;; TYPE: DNA
;; ORGANISM: Homo Sapien
US-10-216-167-39

Query Match 1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAAAAA 1776
Db 1729 AATGAAAAA 1776

RESULT 84
US-10-216-168-39
;; Sequence 39, Application US/10216168
;; Publication No. US20030100712A1
;; GENERAL INFORMATION:
;; APPLICANT: Baker, Kevin P.
;; APPLICANT: Desnoyers, Luc
;; APPLICANT: Gerritsen, Mary
;; APPLICANT: Goddard, Audrey
;; APPLICANT: Godowski, Paul J.
;; APPLICANT: Grimaldi, J. Christopher
;; APPLICANT: Gurney, Austin L.
;; APPLICANT: Smith, Victoria
;; APPLICANT: Stephan, Jean-Philippe F.
;; APPLICANT: Watanabe, Colin L.
;; APPLICANT: Wood, William I.
;; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;; FILE REFERENCE: P3530PIC10
;; CURRENT FILING DATE: 2002-08-09
;; PRIOR APPLICATION NUMBER: 60/059113
;; PRIOR FILING DATE: 1997-09-17
;; PRIOR APPLICATION NUMBER: 60/062287
;; PRIOR FILING DATE: 1997-10-17
;; PRIOR APPLICATION NUMBER: 60/063549
;; PRIOR FILING DATE: 1997-10-28
;; PRIOR APPLICATION NUMBER: 60/064103
;; PRIOR FILING DATE: 1997-10-31
;; PRIOR APPLICATION NUMBER: 60/069873
;; PRIOR FILING DATE: 1997-12-17
;; PRIOR APPLICATION NUMBER: 60/078910
;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/079294
;; PRIOR FILING DATE: 1998-03-25
;; PRIOR APPLICATION NUMBER: 60/079656
;; PRIOR FILING DATE: 1998-03-26
;; PRIOR APPLICATION NUMBER: 60/079728
;; PRIOR FILING DATE: 1998-03-27
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 246
;; SEQ ID NO 39
;; LENGTH: 1837
;; TYPE: DNA
;; ORGANISM: Homo Sapien
US-10-216-168-39

Query Match 1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAA 3762
 |||||
 Db 1729 AATGAA 1776

RESULT 85

US-10-219-065-39
 ; Sequence 39, Application US/10219065
 ; Publication No. US20030100713A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Gerritsen, Mary
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Grimaldi, J. Christopher
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Stephan, Jean-Philippe F.
 ; APPLICANT: Watanabe, Colin L.
 ; APPLICANT: Wood, William I.
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE OF INVENTION: ACIDS ENCODING THE SAME
 ; FILE REFERENCE: P3530PIC39
 ; CURRENT APPLICATION NUMBER: US/10/219,065
 ; CURRENT FILING DATE: 2002-08-13
 ; PRIOR APPLICATION NUMBER: 10/119,480
 ; PRIOR FILING DATE: 2002-04-09
 ; PRIOR APPLICATION NUMBER: 60/059113
 ; PRIOR FILING DATE: 1997-09-17
 ; PRIOR APPLICATION NUMBER: 60/062287
 ; PRIOR FILING DATE: 1997-10-17
 ; PRIOR APPLICATION NUMBER: 60/063549
 ; PRIOR FILING DATE: 1997-10-28
 ; PRIOR APPLICATION NUMBER: 60/064103
 ; PRIOR FILING DATE: 1997-10-31
 ; PRIOR APPLICATION NUMBER: 60/069873
 ; PRIOR FILING DATE: 1997-12-17
 ; PRIOR APPLICATION NUMBER: 60/078910
 ; PRIOR FILING DATE: 1998-03-20
 ; PRIOR APPLICATION NUMBER: 60/079294
 ; PRIOR FILING DATE: 1998-03-25
 ; PRIOR APPLICATION NUMBER: 60/079656
 ; PRIOR FILING DATE: 1998-03-26
 ; PRIOR APPLICATION NUMBER: 60/079728
 ; PRIOR FILING DATE: 1998-03-27
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 246
 ; SEQ ID NO 39
 ; LENGTH: 1837
 ; TYPE: DNA
 ; ORGANISM: Homo Sapien
 US-10-219-065-39

QY 3715 AATGAA 3762
 |||||
 Db 1729 AATGAA 1776

RESULT 86

US-10-219-071-39
 ; Sequence 39, Application US/10219071
 ; Publication No. US20030100714A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Gerritsen, Mary
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Grimaldi, J. Christopher
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Stephan, Jean-Philippe F.
 ; APPLICANT: Watanabe, Colin L.
 ; APPLICANT: Wood, William I.
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE OF INVENTION: ACIDS ENCODING THE SAME
 ; FILE REFERENCE: P3530PIC70
 ; CURRENT APPLICATION NUMBER: US/10/219,074
 ; CURRENT FILING DATE: 2002-08-13
 ; PRIOR APPLICATION NUMBER: 10/119,480
 ; PRIOR FILING DATE: 2002-04-09
 ; PRIOR APPLICATION NUMBER: 60/059113
 ; PRIOR FILING DATE: 1997-09-17
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 246
 ; SEQ ID NO 39
 ; LENGTH: 1837
 ; TYPE: DNA
 ; ORGANISM: Homo Sapien
 US-10-219-065-39

Query Match 1.3%; Score 48; DB 15; Length 1837;
 Best Local Similarity 100.0%; Pred. No. 3.2e-11;
 Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAA 3762
 |||||
 Db 1729 AATGAA 1776

RESULT 87

US-10-219-074-39
 ; Sequence 39, Application US/10219074
 ; Publication No. US20030100715A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Gerritsen, Mary
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Grimaldi, J. Christopher
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Stephan, Jean-Philippe F.
 ; APPLICANT: Watanabe, Colin L.
 ; APPLICANT: Wood, William I.
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE OF INVENTION: ACIDS ENCODING THE SAME
 ; FILE REFERENCE: P3530PIC70
 ; CURRENT APPLICATION NUMBER: US/10/219,074
 ; CURRENT FILING DATE: 2002-08-13
 ; PRIOR APPLICATION NUMBER: 10/119,480
 ; PRIOR FILING DATE: 2002-04-09
 ; PRIOR APPLICATION NUMBER: 60/059113
 ; PRIOR FILING DATE: 1997-09-17

; APPLICANT: Grimaldi, J. Christopher
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Stephan, Jean-Philippe F.
 ; APPLICANT: Watanabe, Colin L.
 ; APPLICANT: Wood, William I.
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE OF INVENTION: ACIDS ENCODING THE SAME
 ; FILE REFERENCE: P3530PIC46
 ; CURRENT APPLICATION NUMBER: US/10/219,071
 ; CURRENT FILING DATE: 2002-08-13
 ; PRIOR APPLICATION NUMBER: 10/119,480
 ; PRIOR FILING DATE: 2002-04-09
 ; PRIOR APPLICATION NUMBER: 60/059113
 ; PRIOR FILING DATE: 1997-09-17
 ; PRIOR APPLICATION NUMBER: 60/062287
 ; PRIOR FILING DATE: 1997-10-17
 ; PRIOR APPLICATION NUMBER: 60/063549
 ; PRIOR FILING DATE: 1997-10-28
 ; PRIOR APPLICATION NUMBER: 60/064103
 ; PRIOR FILING DATE: 1997-10-31
 ; PRIOR APPLICATION NUMBER: 60/069873
 ; PRIOR FILING DATE: 1997-12-17
 ; PRIOR APPLICATION NUMBER: 60/078910
 ; PRIOR FILING DATE: 1998-03-20
 ; PRIOR APPLICATION NUMBER: 60/079294
 ; PRIOR FILING DATE: 1998-03-25
 ; PRIOR APPLICATION NUMBER: 60/079656
 ; PRIOR FILING DATE: 1998-03-26
 ; PRIOR APPLICATION NUMBER: 60/079728
 ; PRIOR FILING DATE: 1998-03-27
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 246
 ; SEQ ID NO 39
 ; LENGTH: 1837
 ; TYPE: DNA
 ; ORGANISM: Homo Sapien
 US-10-219-071-39

Query Match 1.3%; Score 48; DB 15; Length 1837;
 Best Local Similarity 100.0%; Pred. No. 3.2e-11;
 Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAA 3762
 |||||
 Db 1729 AATGAA 1776

RESULT 87

US-10-219-074-39
 ; Sequence 39, Application US/10219074
 ; Publication No. US20030100715A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Gerritsen, Mary
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Grimaldi, J. Christopher
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Stephan, Jean-Philippe F.
 ; APPLICANT: Watanabe, Colin L.
 ; APPLICANT: Wood, William I.
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE OF INVENTION: ACIDS ENCODING THE SAME
 ; FILE REFERENCE: P3530PIC70
 ; CURRENT APPLICATION NUMBER: US/10/219,074
 ; CURRENT FILING DATE: 2002-08-13
 ; PRIOR APPLICATION NUMBER: 10/119,480
 ; PRIOR FILING DATE: 2002-04-09
 ; PRIOR APPLICATION NUMBER: 60/059113
 ; PRIOR FILING DATE: 1997-09-17

Qy	3715	AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	3762	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA			
Matches	48;	Conservative	0;	Mismatches	0;	Indels	0;

; APPLICANT: Gurney, Austin L

;; PRIOR FILING DATE: 1997-10-

```
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 39
; LENGTH: 1837
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-219-471-39

Query Match
Best Local Similarity 1.3%; Score 48; DB 15; Length 1837;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3715 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 1729 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1776

RESULT 93
US-10-219-473-39
; Sequence 39, Application US/10219473
; Publication No. US20030100721A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3530P1C35
; CURRENT APPLICATION NUMBER: US/10/219,473
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: 10/119,480
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 39
; LENGTH: 1837
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-219-471-39

Query Match
Best Local Similarity 1.3%; Score 48; DB 15; Length 1837;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3715 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 1729 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1776

RESULT 94
US-10-219-476-39
; Sequence 39, Application US/10219476
; Publication No. US20030100722A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3530P1C26
; CURRENT APPLICATION NUMBER: US/10/219,476
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: 10/119,480
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 39
; LENGTH: 1837
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-219-476-39

Query Match
Best Local Similarity 1.3%; Score 48; DB 15; Length 1837;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3715 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 1729 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1776
```


;	PRIOR APPLICATION NUMBER:	60/099598
;	PRIOR FILING DATE:	1998-09-09
;	PRIOR APPLICATION NUMBER:	60/099803
;	PRIOR FILING DATE:	1998-09-10
;	PRIOR APPLICATION NUMBER:	60/099811
;	PRIOR FILING DATE:	1998-09-10
;	PRIOR APPLICATION NUMBER:	60/099812
;	PRIOR FILING DATE:	1998-09-10
;	PRIOR APPLICATION NUMBER:	60/099816
;	PRIOR FILING DATE:	1998-09-10
;	PRIOR APPLICATION NUMBER:	60/100038
;	PRIOR FILING DATE:	1998-09-11
;	PRIOR APPLICATION NUMBER:	60/100385
;	PRIOR FILING DATE:	1998-09-15
;	PRIOR APPLICATION NUMBER:	60/100390
;	PRIOR FILING DATE:	1998-09-15
;	PRIOR APPLICATION NUMBER:	60/100627
;	PRIOR FILING DATE:	1998-09-16
;	PRIOR APPLICATION NUMBER:	60/100848
;	PRIOR FILING DATE:	1998-09-18
;	PRIOR APPLICATION NUMBER:	60/100919
;	PRIOR FILING DATE:	1998-09-17
;	PRIOR APPLICATION NUMBER:	60/101477
;	PRIOR FILING DATE:	1998-09-23
;	PRIOR APPLICATION NUMBER:	60/101738
;	PRIOR FILING DATE:	1998-09-24
;	PRIOR APPLICATION NUMBER:	60/101741
;	PRIOR FILING DATE:	1998-09-24
;	PRIOR APPLICATION NUMBER:	60/101786
;	PRIOR FILING DATE:	1998-09-25
;	PRIOR APPLICATION NUMBER:	60/101916
;	PRIOR FILING DATE:	1998-09-24
;	PRIOR APPLICATION NUMBER:	60/101922
;	PRIOR FILING DATE:	1998-09-24
;	PRIOR APPLICATION NUMBER:	60/106178
;	PRIOR FILING DATE:	1998-10-28
;	PRIOR APPLICATION NUMBER:	60/106248
;	PRIOR FILING DATE:	1998-10-29
;	PRIOR APPLICATION NUMBER:	60/106464
;	PRIOR FILING DATE:	1998-10-30
;	PRIOR APPLICATION NUMBER:	60/106905
;	PRIOR FILING DATE:	1998-11-03
;	PRIOR APPLICATION NUMBER:	60/108787
;	PRIOR FILING DATE:	1998-11-17
;	PRIOR APPLICATION NUMBER:	60/108801
;	PRIOR FILING DATE:	1998-11-17
;	PRIOR APPLICATION NUMBER:	60/108849
;	PRIOR FILING DATE:	1998-11-18
;	PRIOR APPLICATION NUMBER:	60/112422
;	PRIOR FILING DATE:	1998-12-15
;	PRIOR APPLICATION NUMBER:	60/113296
;	PRIOR FILING DATE:	1998-12-22
;	PRIOR APPLICATION NUMBER:	60/113605
;	PRIOR FILING DATE:	1998-12-23
;	PRIOR APPLICATION NUMBER:	60/113621
;	PRIOR FILING DATE:	1998-12-23
;	PRIOR APPLICATION NUMBER:	60/115558
;	PRIOR FILING DATE:	1999-01-12
;	PRIOR APPLICATION NUMBER:	60/115565
;	PRIOR FILING DATE:	1999-01-12
;	PRIOR APPLICATION NUMBER:	60/115733
;	PRIOR FILING DATE:	1999-01-12
;	PRIOR APPLICATION NUMBER:	60/119549
;	PRIOR FILING DATE:	1999-02-10
;	PRIOR APPLICATION NUMBER:	60/123618
;	PRIOR FILING DATE:	1999-03-10
;	PRIOR APPLICATION NUMBER:	60/125259
;	PRIOR FILING DATE:	1999-03-19
;	PRIOR APPLICATION NUMBER:	60/125775
;	PRIOR FILING DATE:	1999-03-23
;	PRIOR APPLICATION NUMBER:	60/126773
;	PRIOR FILING DATE:	1999-03-29
;	PRIOR APPLICATION NUMBER:	60/127887

1 PRIOR FILING DATE: 1999-04-05
 2 PRIOR APPLICATION NUMBER: 60/130232
 3 PRIOR FILING DATE: 1999-04-21
 4 PRIOR APPLICATION NUMBER: 60/131022
 5 PRIOR FILING DATE: 1999-04-26
 6 PRIOR APPLICATION NUMBER: 60/131270
 7 PRIOR FILING DATE: 1999-04-27
 8 PRIOR APPLICATION NUMBER: 60/131291
 9 PRIOR FILING DATE: 1999-04-27
 10 PRIOR APPLICATION NUMBER: 60/131445
 11 PRIOR FILING DATE: 1999-04-28
 12 PRIOR APPLICATION NUMBER: 60/134287
 13 PRIOR FILING DATE: 1999-05-14
 14 PRIOR APPLICATION NUMBER: 60/140650
 15 PRIOR FILING DATE: 1999-06-22
 16 PRIOR APPLICATION NUMBER: 60/140723
 17 PRIOR FILING DATE: 1999-06-22
 18 PRIOR APPLICATION NUMBER: 60/141037
 19 PRIOR FILING DATE: 1999-06-23
 20 PRIOR APPLICATION NUMBER: 60/144758
 21 PRIOR FILING DATE: 1999-07-20
 22 PRIOR APPLICATION NUMBER: 60/145698
 23 PRIOR FILING DATE: 1999-07-26
 24 PRIOR APPLICATION NUMBER: 60/146222
 25 PRIOR FILING DATE: 1999-07-28
 26 PRIOR APPLICATION NUMBER: 60/146963
 27 PRIOR FILING DATE: 1999-08-03
 28 PRIOR APPLICATION NUMBER: 60/149320
 29 PRIOR FILING DATE: 1999-08-17
 30 PRIOR APPLICATION NUMBER: 60/149638
 31 PRIOR FILING DATE: 1999-08-17
 32 PRIOR APPLICATION NUMBER: 60/151733
 33 PRIOR FILING DATE: 1999-08-31
 34 PRIOR APPLICATION NUMBER: 60/164418
 35 PRIOR FILING DATE: 1999-11-09
 36 PRIOR APPLICATION NUMBER: 60/166361
 37 PRIOR FILING DATE: 1999-11-16
 38 PRIOR APPLICATION NUMBER: 60/169445
 39 PRIOR FILING DATE: 1999-12-07
 40 PRIOR APPLICATION NUMBER: 60/169495
 41 PRIOR FILING DATE: 1999-12-07
 42 PRIOR APPLICATION NUMBER: 60/169835

RESULT 97
US-10-227-876-39
; Sequence 39, Application US/10227876
; Publication No. US20030100725A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3530PIC73
; CURRENT APPLICATION NUMBER: US/10/227,876
; CURRENT FILING DATE: 2002-08-26

, PRIOR APPLICATION NUMBER: 10/119,480
, PRIOR FILING DATE: 2002-04-09
, PRIOR APPLICATION NUMBER: 60/059113
, PRIOR FILING DATE: 1997-09-17
, PRIOR APPLICATION NUMBER: 60/062287
, PRIOR FILING DATE: 1997-10-17
, PRIOR APPLICATION NUMBER: 60/063549
, PRIOR FILING DATE: 1997-10-28
, PRIOR APPLICATION NUMBER: 60/064103
, PRIOR FILING DATE: 1997-10-31
, PRIOR APPLICATION NUMBER: 60/069873
, PRIOR FILING DATE: 1997-12-17
, PRIOR APPLICATION NUMBER: 60/078910
, PRIOR FILING DATE: 1998-03-20
, PRIOR APPLICATION NUMBER: 60/079294
, PRIOR FILING DATE: 1998-03-25
, PRIOR APPLICATION NUMBER: 60/079656
, PRIOR FILING DATE: 1998-03-26
, PRIOR APPLICATION NUMBER: 60/079728
, PRIOR FILING DATE: 1998-03-27
, PRIOR APPLICATION NUMBER: 60/081819
, PRIOR FILING DATE: 1998-04-15
, PRIOR APPLICATION NUMBER: 60/081955
, PRIOR FILING DATE: 1998-04-15
, PRIOR APPLICATION NUMBER: 60/082804
, PRIOR FILING DATE: 1998-04-22
, PRIOR APPLICATION NUMBER: 60/084441
, PRIOR FILING DATE: 1998-05-06
, PRIOR APPLICATION NUMBER: 60/085323
, PRIOR FILING DATE: 1998-05-13
, PRIOR APPLICATION NUMBER: 60/085579
, PRIOR FILING DATE: 1998-05-15
, PRIOR APPLICATION NUMBER: 60/086392
, PRIOR FILING DATE: 1998-05-22
, PRIOR APPLICATION NUMBER: 60/089532
, PRIOR FILING DATE: 1998-06-17
, PRIOR APPLICATION NUMBER: 60/089538
, PRIOR FILING DATE: 1998-06-17
, PRIOR APPLICATION NUMBER: 60/089905
, PRIOR FILING DATE: 1998-06-18
, PRIOR APPLICATION NUMBER: 60/090472
, PRIOR FILING DATE: 1998-06-24
, PRIOR APPLICATION NUMBER: 60/090557
, PRIOR FILING DATE: 1998-06-24
, PRIOR APPLICATION NUMBER: 60/090691
, PRIOR FILING DATE: 1998-06-25
, PRIOR APPLICATION NUMBER: 60/090695
, PRIOR FILING DATE: 1998-06-25
, PRIOR APPLICATION NUMBER: 60/091982
, PRIOR FILING DATE: 1998-07-07
, PRIOR APPLICATION NUMBER: 60/095302
, PRIOR FILING DATE: 1998-08-04
, PRIOR APPLICATION NUMBER: 60/095318
, PRIOR FILING DATE: 1998-08-04
, PRIOR APPLICATION NUMBER: 60/095916
, PRIOR FILING DATE: 1998-08-10
, PRIOR APPLICATION NUMBER: 60/096146
, PRIOR FILING DATE: 1998-08-11
, PRIOR APPLICATION NUMBER: 60/096791
, PRIOR FILING DATE: 1998-08-17
, PRIOR APPLICATION NUMBER: 60/097986
, PRIOR FILING DATE: 1998-08-25
, PRIOR APPLICATION NUMBER: 60/098544
, PRIOR FILING DATE: 1998-08-31
, PRIOR APPLICATION NUMBER: 60/099596
, PRIOR FILING DATE: 1998-09-09
, PRIOR APPLICATION NUMBER: 60/099598
, PRIOR FILING DATE: 1998-09-09
, PRIOR APPLICATION NUMBER: 60/099803
, PRIOR FILING DATE: 1998-09-10
, PRIOR APPLICATION NUMBER: 60/099811
, PRIOR FILING DATE: 1998-09-10
, PRIOR APPLICATION NUMBER: 60/099812
, PRIOR FILING DATE: 1998-09-10
, PRIOR APPLICATION NUMBER: 60/099816
, PRIOR FILING DATE: 1998-09-10
, PRIOR APPLICATION NUMBER: 60/100038
, PRIOR FILING DATE: 1998-09-11
, PRIOR APPLICATION NUMBER: 60/100385
, PRIOR FILING DATE: 1998-09-15
, PRIOR APPLICATION NUMBER: 60/100390
, PRIOR FILING DATE: 1998-09-15
, PRIOR APPLICATION NUMBER: 60/100627
, PRIOR FILING DATE: 1998-09-16
, PRIOR APPLICATION NUMBER: 60/100848
, PRIOR FILING DATE: 1998-09-16
, PRIOR APPLICATION NUMBER: 60/100919
, PRIOR FILING DATE: 1998-09-17
, PRIOR APPLICATION NUMBER: 60/101477
, PRIOR FILING DATE: 1998-09-23
, PRIOR APPLICATION NUMBER: 60/101738
, PRIOR FILING DATE: 1998-09-24
, PRIOR APPLICATION NUMBER: 60/101741
, PRIOR FILING DATE: 1998-09-24
, PRIOR APPLICATION NUMBER: 60/101786
, PRIOR FILING DATE: 1998-09-25
, PRIOR APPLICATION NUMBER: 60/101916
, PRIOR FILING DATE: 1998-09-24
, PRIOR APPLICATION NUMBER: 60/101922
, PRIOR FILING DATE: 1998-09-24
, PRIOR APPLICATION NUMBER: 60/106178
, PRIOR FILING DATE: 1998-10-28
, PRIOR APPLICATION NUMBER: 60/106248
, PRIOR FILING DATE: 1998-10-29
, PRIOR APPLICATION NUMBER: 60/106464
, PRIOR FILING DATE: 1998-10-30
, PRIOR APPLICATION NUMBER: 60/106905
, PRIOR FILING DATE: 1998-11-03
, PRIOR APPLICATION NUMBER: 60/108787
, PRIOR FILING DATE: 1998-11-17
, PRIOR APPLICATION NUMBER: 60/108801
, PRIOR FILING DATE: 1998-11-17
, PRIOR APPLICATION NUMBER: 60/108849
, PRIOR FILING DATE: 1998-11-18
, PRIOR APPLICATION NUMBER: 60/112422
, PRIOR FILING DATE: 1998-12-15
, PRIOR APPLICATION NUMBER: 60/113296
, PRIOR FILING DATE: 1998-12-22
, PRIOR APPLICATION NUMBER: 60/113605
, PRIOR FILING DATE: 1998-12-23
, PRIOR APPLICATION NUMBER: 60/113621
, PRIOR FILING DATE: 1998-12-23
, PRIOR APPLICATION NUMBER: 60/115558
, PRIOR FILING DATE: 1999-01-12
, PRIOR APPLICATION NUMBER: 60/115565
, PRIOR FILING DATE: 1999-01-12
, PRIOR APPLICATION NUMBER: 60/115733
, PRIOR FILING DATE: 1999-01-12
, PRIOR APPLICATION NUMBER: 60/119549
, PRIOR FILING DATE: 1999-02-10
, PRIOR APPLICATION NUMBER: 60/123618
, PRIOR FILING DATE: 1999-03-10
, PRIOR APPLICATION NUMBER: 60/125259
, PRIOR FILING DATE: 1999-03-19
, PRIOR APPLICATION NUMBER: 60/125775
, PRIOR FILING DATE: 1999-03-23
, PRIOR APPLICATION NUMBER: 60/126773
, PRIOR FILING DATE: 1999-03-29
, PRIOR APPLICATION NUMBER: 60/127887
, PRIOR FILING DATE: 1999-04-05
, PRIOR APPLICATION NUMBER: 60/130232
, PRIOR FILING DATE: 1999-04-21
, PRIOR APPLICATION NUMBER: 60/131022
, PRIOR FILING DATE: 1999-04-26
, PRIOR APPLICATION NUMBER: 60/131270
, PRIOR FILING DATE: 1999-04-27

[illegible]

; PRIOR APPLICATION NUMBER: 60/100390
; PRIOR FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: 60/100627
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: 60/100848
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/100919
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: 60/101477
; PRIOR FILING DATE: 1998-09-23
; PRIOR APPLICATION NUMBER: 60/101738
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/101741
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/101786
; PRIOR FILING DATE: 1998-09-25
; PRIOR APPLICATION NUMBER: 60/101916
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/101922
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/106178
; PRIOR FILING DATE: 1998-10-28
; PRIOR APPLICATION NUMBER: 60/106248
; PRIOR FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: 60/106464
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: 60/106905
; PRIOR FILING DATE: 1998-11-03
; PRIOR APPLICATION NUMBER: 60/108787
; PRIOR FILING DATE: 1998-11-17
; PRIOR APPLICATION NUMBER: 60/108801
; PRIOR FILING DATE: 1998-11-17
; PRIOR APPLICATION NUMBER: 60/108849
; PRIOR FILING DATE: 1998-11-18
; PRIOR APPLICATION NUMBER: 60/112422
; PRIOR FILING DATE: 1998-12-15
; PRIOR APPLICATION NUMBER: 60/113296
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: 60/113605
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: 60/113621
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: 60/115558
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/115565
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/115733
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/119549
; PRIOR FILING DATE: 1999-02-10
; PRIOR APPLICATION NUMBER: 60/123618
; PRIOR FILING DATE: 1999-03-10
; PRIOR APPLICATION NUMBER: 60/125259
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 60/125775
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: 60/126773
; PRIOR FILING DATE: 1999-03-29
; PRIOR APPLICATION NUMBER: 60/127887
; PRIOR FILING DATE: 1999-04-05
; PRIOR APPLICATION NUMBER: 60/130232
; PRIOR FILING DATE: 1999-04-21
; PRIOR APPLICATION NUMBER: 60/131022
; PRIOR FILING DATE: 1999-04-26
; PRIOR APPLICATION NUMBER: 60/131270
; PRIOR FILING DATE: 1999-04-27
; PRIOR APPLICATION NUMBER: 60/131291
; PRIOR FILING DATE: 1999-04-27
; PRIOR APPLICATION NUMBER: 60/131445
; PRIOR FILING DATE: 1999-04-28
; PRIOR APPLICATION NUMBER: 60/134287
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/140650

; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: 60/140723
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: 60/141037
; PRIOR FILING DATE: 1999-06-23
; PRIOR APPLICATION NUMBER: 60/144758
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 60/145698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: 60/146222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: 60/146963
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: 60/149320
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: 60/149638
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: 60/151733
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: 60/164418
; PRIOR FILING DATE: 1999-11-09
; PRIOR APPLICATION NUMBER: 60/166361
; PRIOR FILING DATE: 1999-11-16
; PRIOR APPLICATION NUMBER: 60/169445
; PRIOR FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: 60/169495
; PRIOR FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: 60/169835

Query Match 1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
|||
Db 1729 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1776

RESULT 99
US-10-229-974-39
; Sequence 39, Application US/10229974
; Publication No. US20030100727A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3530PIC91
; CURRENT APPLICATION NUMBER: US/10/229,974
; CURRENT FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: 10/119,480
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20

; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 39
; LENGTH: 1837
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-229-974-39

Query Match 1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred.No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3715 AATGAA 3762
DB 1729 AATGAA 1776

RESULT 100

US-10-230-024-39
; Sequence 39, Application US/10230024
; Publication No. US20030100728A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Destoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: F3530P1C89
; CURRENT APPLICATION NUMBER: US/10/230,024
; CURRENT FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: 10/119,480
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 39
; LENGTH: 1837
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-230-024-39

Query Match

1.3%; Score 48; DB 15; Length 1837;

Best Local Similarity 100.0%; Pred.No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3715 AATGAA 3762
DB 1729 AATGAA 1776

Search completed: April 23, 2004, 14:38:17
Job time : 1475 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 23, 2004, 03:08:45 ; Search time 1339 seconds
(without alignments)

11935.565 Million cell updates/sec

Title: US-10-005-907-1

Perfect score: 3762

Sequence: 1 gagaaacggctcactgtga.....aaaaaaaaaaaaaaaaaaaaa 3762

Scoring table: OLIGO_NUC

Gapop_60.0, Gapext 60.0

Searched: 3373863 seqs, 2124099041 residues

Word size: 0

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 150 summaries

Database: N Geneseq_29Jan04.*
1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002as.*
7: Geneseqn2002bs.*
8: Geneseqn2003as.*
9: Geneseqn2003bs.*
10: Geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3762	100.0	3762	6	ABN81319 Human mas
2	733	19.5	1128	5	ABV20298 Human pro
3	733	19.5	1128	5	ABV26130 Human pro
4	733	19.5	1128	5	ABV26458 Human pro
5	733	19.5	1128	5	ABV20617 Human pro
6	733	19.5	1128	5	ABV20806 Human pro
7	733	19.5	1128	5	ABV20439 Human pro
8	733	19.5	1128	5	ABV26275 Human pro
9	733	19.5	1128	5	ABV26523 Human pro
10	506	13.5	667	9	ADC30760 Human nov
11	344	9.1	397	5	ABV10439 Human pro
12	344	9.1	433	5	ABV01270 Human pro
13	342	9.1	437	5	ABV31609 Human pro
14	342	9.1	437	5	ABV40245 Human pro
15	342	9.1	437	5	ABV40577 Human pro
16	244	6.5	446	5	ABV00935 Human pro
17	213	5.7	404	5	ABV10104 Human pro
18	213	5.7	421	5	ABV31276 Human pro
19	204	5.4	359	5	ABV09593 Human pro
20	203	5.4	400	5	ABV00424 Human pro
21	203	5.4	401	5	ABV39737 Human pro
22	203	5.4	401	5	ABV30769 Human pro
23	203	5.4	401	5	ABV39958 Human pro

24	62	1.6	296	5	ABV30990	Human pro
25	62	1.6	337	5	ABV00647	Human pro
26	62	1.6	425	5	ABV09816	Human pro
27	52	1.4	381	5	ABV56428	Human pro
28	50	1.3	345	4	AAI90783	Human pol
29	50	1.3	424	5	ABV57393	Human pro
30	50	1.3	449	4	AAI83050	Human pol
31	50	1.3	1702	3	AAF15960	Human pro
32	50	1.3	1820	9	ADD18824	Human dis
33	50	1.3	3602	7	ACF12861	Human cer
34	50	1.3	3607	2	AAH87412	Hepatocel
35	49	1.3	265	5	ABV48235	Human pro
36	49	1.3	287	5	ABV18451	Human pro
37	49	1.3	381	4	AAI84445	Human pol
38	49	1.3	476	6	ABZ78073	Human bre
39	49	1.3	497	5	ABV58273	Human pro
40	49	1.3	882	2	AAI72173	Alzheimer
41	49	1.3	903	3	AAI79727	Human sec
42	49	1.3	1057	4	AAH33675	Human col
43	49	1.3	1263	3	AAI79719	Human sec
44	49	1.3	1373	3	AAI77616	Human can
45	49	1.3	1429	6	ABV94043	Breast ca
46	49	1.3	1676	4	AAH72649	Human cer
47	49	1.3	1702	5	ABV27467	Human pro
48	49	1.3	1702	5	ABV27724	Human pro
49	49	1.3	1702	5	ABV25281	Human pro
50	49	1.3	1702	5	ABV21893	Human pro
51	49	1.3	1933	4	AAI007369	Human DNA
52	49	1.3	1933	6	ABK92253	Prostate
53	49	1.3	1998	4	AD16758	Human nov
54	49	1.3	1998	9	ADC22050	Human cDN
55	49	1.3	2038	7	ACD13422	Human DNA
56	49	1.3	2106	7	ABQ71965	Human cDN
57	49	1.3	2143	6	ABQ54860	Human ova
58	49	1.3	2543	3	AAI76644	Human ORF
59	49	1.3	2713	6	AAI43556	Human CD2
60	49	1.3	2713	6	AAI43566	Human CD2
61	49	1.3	2713	6	AAI43565	Human CD2
62	49	1.3	2713	6	AAI43564	Human CD2
63	49	1.3	2713	6	AAI43567	Human CD2
64	49	1.3	2851	6	ABN86559	Canine lo
65	49	1.3	2851	6	ABN86560	Canine CD
66	49	1.3	2944	7	ABT18794	Aspergill
67	49	1.3	3362	7	ABT20610	Aspergill
68	49	1.3	4326	6	ABQ92014	Human pol
69	49	1.3	4344	7	ABT18200	Aspergill
70	49	1.3	5362	7	ABT20014	Aspergill
71	48	1.3	295	4	AAH69364	Human cer
72	48	1.3	337	7	ABX40871	Bovine BS
73	48	1.3	401	8	ACH29107	Human adu
74	48	1.3	404	5	ABV44841	Human pro
75	48	1.3	418	4	AAI87343	Human pol
76	48	1.3	419	4	AAI88543	Human pol
77	48	1.3	546	4	AAH71167	Human cer
78	48	1.3	619	3	AAI26393	Human sec
79	48	1.3	865	6	ABN98811	Arabidops
80	48	1.3	903	4	AAI87708	Human pol
81	48	1.3	1136	3	AAI93117	Human sec
82	48	1.3	1467	3	AAI79042	Human sec
83	48	1.3	1510	6	ABK87557	CDNA enco
84	48	1.3	1544	3	AAI46483	CDNA enco
85	48	1.3	1837	6	ABK33555	CDNA enco
86	48	1.3	1837	7	ACA66860	CDNA enco
87	48	1.3	1837	7	ACD68612	Novel hum
88	48	1.3	1837	7	ACA68516	Novel hum
89	48	1.3	1837	8	ABT44245	Human PRO
90	48	1.3	1837	8	ABT44528	Human PRO
91	48	1.3	1837	8	ABT433901	Human men
92	48	1.3	1837	8	ABD83529	Novel hum
93	48	1.3	1837	8	ABD80635	Novel hum
94	48	1.3	1837	8	ABD73176	Novel hum
95	48	1.3	1837	8	ABD78258	Novel hum
96	48	1.3	1837	8	ABD78258	Novel hum

97 48 1.3 1837 9 ADB84906 Human PRO
 98 48 1.3 1837 9 ADB78012 Novel hum
 99 48 1.3 1837 9 ADB87078 Human PRO
 100 48 1.3 1837 9 ADB84660 Human PRO
 101 48 1.3 1837 9 ADB83775 Novel hum
 102 48 1.3 1837 9 ADB72930 Novel hum
 103 48 1.3 1837 9 ADC36768 Human PRO
 104 48 1.3 1837 9 ADC21758 Human PRO
 105 48 1.3 1837 9 ADC49789 Novel hum
 106 48 1.3 1837 9 ADC48988 Novel hum
 107 48 1.3 1837 9 ADC49505 Novel hum
 108 48 1.3 1837 9 ADC47366 Novel hum
 109 48 1.3 1837 9 ADC47111 Novel hum
 110 48 1.3 1837 9 ADC77986 Novel hum
 111 48 1.3 1837 9 ADD06221 Novel hum
 112 48 1.3 1837 9 ADC77740 Novel hum
 113 48 1.3 1837 9 ADD50703 Novel hum
 114 48 1.3 1837 9 ADD50949 Novel hum
 115 48 1.3 1837 9 ADD50430 Human PRO
 116 48 1.3 1837 9 ADD50184 Human PRO
 117 48 1.3 1837 9 ADD51195 Novel hum
 118 48 1.3 1837 10 ADC48742 Novel hum
 119 48 1.3 1837 10 ADE20913 Novel hum
 120 48 1.3 1837 10 ADE05757 Human PRO
 121 48 1.3 1837 10 ADD74986 Human PRO
 122 48 1.3 1837 10 ADD75732 Novel hum
 123 48 1.3 1837 10 ADD4964 Novel hum
 124 48 1.3 1837 10 ADD86790 Novel hum
 125 48 1.3 1837 10 ADE20667 Novel hum
 126 48 1.3 1837 10 ADE38964 Novel hum
 127 48 1.3 1837 10 ADE05511 Human PRO
 128 48 1.3 1837 10 ADD73496 Human PRO
 129 48 1.3 1837 10 ADD78336 Novel hum
 130 48 1.3 1837 10 ADE21159 Novel hum
 131 48 1.3 1837 10 ADE20421 Novel hum
 132 48 1.3 1837 10 ADE20421 Novel hum
 133 48 1.3 1837 10 ADD75486 Human PRO
 134 48 1.3 1837 10 ADD74002 Human PRO
 135 48 1.3 1837 10 ADD74248 Human PRO
 136 48 1.3 1837 10 ADD75978 Novel hum
 137 48 1.3 1837 10 ADD85470 Novel hum
 138 48 1.3 1837 10 ADE05019 Human PRO
 139 48 1.3 1837 10 ADD75232 Human PRO
 140 48 1.3 1837 10 ADD76776 Novel hum
 141 48 1.3 1837 10 ADD86544 Novel hum
 142 48 1.3 1837 10 ADD78012 Novel hum
 143 48 1.3 1837 10 ADD77520 Novel hum
 144 48 1.3 1837 10 ADD77766 Novel hum
 145 48 1.3 1837 10 ADD85224 Novel hum
 146 48 1.3 1837 10 ADD73756 Human PRO
 147 48 1.3 1837 10 ADD74494 Human PRO
 148 48 1.3 1837 10 ADD77022 Novel hum
 149 48 1.3 1837 10 ADD85716 Novel hum
 150 48 1.3 1837 10 ADE05265 Human PRO

ALIGNMENTS

RESULT 1
 ID AEN81319 standard; cDNA; 3762 BP.
 XX
 AC AEN81319;
 XX
 DT 30-AUG-2002 (first entry)
 XX
 DE Human mast cell related gene MCL SEQ ID NO 1.
 XX
 KW Human; mast cell; MC; antiallergic; antiinflammatory; antiasthmatic;
 KW vasotrophic; dermatological; allergic; hypersensitivity; rhinitis; asthma;
 KW gene; ss.
 XX

OS Homo sapiens.
 XX Key Location/Qualifiers
 PH CDS 25..432
 FT /*tag= a
 FT /product= "MCL"
 XX
 PN WC2002046389-A2.
 XX
 PD 13-JUN-2002.
 XX
 PD 07-DEC-2001; 2001WO-US046180.
 PF
 PR 08-DEC-2000; 2000US-0251835P.
 PR 14-MAR-2001; 2001US-0275479P.
 PR 28-MAR-2001; 2001US-0279115P.
 PR 02-APR-2001; 2001US-0280143P.
 XX
 PA (UNIO) UCB SA.
 XX
 PI Nocka K, Pirozzi G, Einstein R;
 XX WPI; 2002-508560/54.
 DR P-PSDB; ABB77569.
 DR
 XX
 PT Novel isolated nucleic acids that are differentially expressed in mast
 PT cells in patients with allergic hypersensitivity, encoding proteins
 PT associated with mast cell regranulation and allergic hypersensitivity.
 XX
 PS Claim 1; Page 95-97; 119pp; English.
 XX
 CC The invention relates to isolated nucleic acid (AEN81319-AEN81324),
 CC corresponding to genes differentially expressed in mast cells following
 CC activation or in patients with allergic hypersensitivity disease, (I)
 CC that encodes proteins (ABB77569-ABB77575) (II) or a protein fragment of
 CC (II) if at least 6 amino acids. (II) is useful for identifying binding
 CC partners. (I) or (II) is useful for diagnosing or treating a disease
 CC state (e.g. allergic hypersensitivity, seasonal rhinitis, asthma,
 CC urticaria or atopic dermatitis or mastocytosis) in a subject which
 CC involves determining the level of expression of (I) or (II). A computer
 CC system, comprising a database containing information identifying the
 CC expression level in a tissue or at least one mast cell of (I), is useful
 CC for presenting information to identify the relative expression level of
 CC (I). (II) is used as a marker to detect, diagnose or identify an allergic
 CC response in a patient. The protein can also serve as a target that
 CC modulate gene expression or activity and as an antigen to raise
 CC polyclonal or monoclonal antibodies. (II) is useful for identifying
 CC agents that modulate expression of the protein or agents, such as
 CC agonists or antagonists. The agonists or antagonists are useful for
 CC modulating biological activity and function of (II) and thus are useful
 CC for alleviating disease conditions such as allergic hypersensitivity,
 CC seasonal rhinitis, asthma, urticaria, atopic dermatitis or mastocytosis
 XX
 SQ Sequence 3762 BP; 1220 A; 672 C; 680 G; 1190 T; 0 U; 0 Other;
 Query Match 100.0%; Score 3762; DB 6; Length 3762;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 3762; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GAGAAACCGAGTCACTGTGAAAGATGGAAATATCTCTCGGAAACTCAGTTGCTCG 60
 DB 1 GAGAAACCGAGTCACTGTGAAAGATGGAAATATCTCTCGGAAACTCAGTTGCTCG 60
 QY 61 GGAGAGATCAAAAGAACCCCAAGAAAGAAACCCAGATGAGGAAAGAAACCGCAGGAA 120
 DB 61 GGAGAGATCAAAAGAACCCCAAGAAAGAAACCCAGATGAGGAAAGAAACCGCAGGAA 120
 QY 121 ATGACTACATTGAAAGAAACTCAAGATCAAGTAGAAGAAAGCAAGAGAGTTTCATCC 180
 DB 121 ATGACTACATTGAAAGAAACTCAAGATCAAGTAGAAGAAAGCAAGAGAGTTTCATCC 180
 QY 181 ACTTCTAATCAGGAAACGAGAAATGGCAGTGGTCTCTGAAGAGTGTGTACACTGTCTATT 240

Db 181 ACTTCTAATCAGGAAACGAGAAATGCGAGTGGTTCTGAAGAGGTGCTACACTGTCAAT 240
Qy 241 AATCACATCCCCCATCAGAGATCTCCCTGAGCTCCAAATGATCATGCTATGAGACATT 300
Db 241 AATCACATCCCCCATCAGAGATCTCCCTGAGCTCCAAATGATCATGCTATGAGACATT 300
Qy 301 GACTCCCTCAAGGAAAGTGAGACAGTTTATGAGAAAGGTGAGAGACAGAAATATGCCCTT 360
Db 301 GACTCCCTCAAGGAAAGTGAGACAGTTTATGAGAAAGGTGAGAGACAGAAATATGCCCTT 360
Qy 361 CTTAGGACTTCTGTAGTAGCCCTGTTCTGCTGACCCCATGAGCATGATGAGAGTTG 420
Db 361 CTTAGGACTTCTGTAGTAGCCCTGTTCTGCTGACCCCATGAGCATGATGAGAGTTG 420
Qy 421 TTTCCACACATAAATCCTCAAGCTGCTTTATCACTTTCCAGCAATGAAGACATGAGAA 480
Db 421 TTTCCACACATAAATCCTCAAGCTGCTTTATCACTTTCCAGCAATGAAGACATGAGAA 480
Qy 481 TAGCAGACTCTGGGAAAGTTTCACTGAGCAGTGCATGAACAATCTCTTCTGCTA 540
Db 481 TAGCAGACTCTGGGAAAGTTTCACTGAGCAGTGCATGAACAATCTCTTCTGCTA 540
Qy 541 AAGTTTGAATAATATCTTATATATATCTTATGAGCACTCTGATATGAGCATCTCTG 600
Db 541 AAGTTTGAATAATATCTTATATATATCTTATGAGCACTCTGATATGAGCATCTCTG 600
Qy 601 TGGCTTAGGTGAATCATAGAAATGACACAATGACCTAAATATTTCTATGTTTTG 660
Db 601 TGGCTTAGGTGAATCATAGAAATGACACAATGACCTAAATATTTCTATGTTTTG 660
Qy 661 TTGTAAGTTTGAAGCATGGAGTGATATAAATAAATCTTTCTAGGACATATATGTA 720
Db 661 TTGTAAGTTTGAAGCATGGAGTGATATAAATAAATCTTTCTAGGACATATATGTA 720
Qy 721 TGAATAATAATTTCTAATCCCTGACTAATGAGACCTCTCTCTAGGCCAAAGAGA 780
Db 721 TGAATAATAATTTCTAATCCCTGACTAATGAGACCTCTCTCTAGGCCAAAGAGA 780
Qy 781 CCTCAGATGACCTGAAGACTGAAATCTGGCCATGATAGAGGAGGTGAGACAC 840
Db 781 CCTCAGATGACCTGAAGACTGAAATCTGGCCATGATAGAGGAGGTGAGACAC 840
Qy 841 TTGTTATACCCCTTCCCTTTTGGAGTTTATGCAAGTGACAGGAGTGAATAGACT 900
Db 841 TTGTTATACCCCTTCCCTTTTGGAGTTTATGCAAGTGACAGGAGTGAATAGACT 900
Qy 901 GATGAAATAGACTGATTTGGCAATAAGAGTCCCAATTCCAACTGACTGCTGTAGAT 960
Db 901 GATGAAATAGACTGATTTGGCAATAAGAGTCCCAATTCCAACTGACTGCTGTAGAT 960
Qy 961 CACACACTGCTGAGGATTCATCTATGAGACTTTGCTACATAACAGAGACCTTGGTT 1020
Db 961 CACACACTGCTGAGGATTCATCTATGAGACTTTGCTACATAACAGAGACCTTGGTT 1020
Qy 1021 TCCCAACCCCTTTATTTAGCTAAAGCATCTTTTCTACTGACTTTAAAGCTTTTGA 1080
Db 1021 TCCCAACCCCTTTATTTAGCTAAAGCATCTTTTCTACTGACTTTTAAAGCTTTTGA 1080
Qy 1081 CAAGCTTAACTCTTTCAACCAATGGCAATCAGACAACTTTTGAATCTACCTATGACCT 1140
Db 1081 CAAGCTTAACTCTTTCAACCAATGGCAATCAGACAACTTTTGAATCTACCTATGACCT 1140
Qy 1141 GTAAGCTCTCTCTGCTTCAAGATCTTGGCTCTTTAAGCTGAACCGATGTCACCTTTCCA 1200
Db 1141 GTAAGCTCTCTCTGCTTCAAGATCTTGGCTCTTTAAGCTGAACCGATGTCACCTTTCCA 1200
Qy 1201 TTTTAACTGATTTAGCTTTGCTTGAATCTGCTCCCTGCTCCCTGCTTAAATGATTAAG 1260
Db 1201 TTTTAACTGATTTAGCTTTGCTTGAATCTGCTCCCTGCTCCCTGCTTAAATGATTAAG 1260
Qy 1261 GTGACCTGACCACTCAGGACACTTTTCTCAGGACCTCTCAGAGTGTATCCAGGCCAT 1320
Db 1261 GTGACCTGACCACTCAGGACACTTTTCTCAGGACCTCTCAGAGTGTATCCAGGCCAT 1320

Qy 1321 GGTAACTCATCTTGGCTCAGAAATCAACTCTTTTAAATATTTTACAGAAATTTGGTTTGG 1380
Db 1321 GGTAACTCATCTTGGCTCAGAAATCAACTCTTTTAAATATTTTACAGAAATTTGGTTTGG 1380
Qy 1381 TTACCAATAAGTCTCCCAAAATATATGTCGAAGAAATCTTCAATTTCCAGGCTGTCACCA 1440
Db 1381 TTACCAATAAGTCTCCCAAAATATATGTCGAAGAAATCTTCAATTTCCAGGCTGTCACCA 1440
Qy 1441 AATTTCAATGCCAACATCTCCCATCAATTTCACTATTTTCACTTTTGGAGTGTAACTA 1500
Db 1441 AATTTCAATGCCAACATCTCCCATCAATTTCACTATTTTCACTTTTGGAGTGTAACTA 1500
Qy 1501 CTCATAAATCTGTGAAGCAAGTGAAGCAAGTGAAGCAAGTGAAGCAAGTGAAGCAAGT 1560
Db 1501 CTCATAAATCTGTGAAGCAAGTGAAGCAAGTGAAGCAAGTGAAGCAAGTGAAGCAAGT 1560
Qy 1561 TTTTCTTTTCTATGCTAGTGAATTTTTCATATAAATTTTTCATATAAATTTTTCATATAA 1620
Db 1561 TTTTCTTTTCTATGCTAGTGAATTTTTCATATAAATTTTTCATATAAATTTTTCATATAA 1620
Qy 1621 TAAATAGTTTTCGACATTTGCTGAGAGAGAAATTTGAAAGTGTCAAAAATAAAAA 1680
Db 1621 TAAATAGTTTTCGACATTTGCTGAGAGAGAAATTTGAAAGTGTCAAAAATAAAAA 1680
Qy 1681 AAGTGAATGAAGCATATATAATTTTCAATTTTTCATATAAATTTTTCATATAAATTTT 1740
Db 1681 AAGTGAATGAAGCATATATAATTTTCAATTTTTCATATAAATTTTTCATATAAATTTT 1740
Qy 1741 GAAGGATTTCTGTTCAATTTAGTAAATTTGAAATTTGAAATTTGAAATTTGAAATTT 1800
Db 1741 GAAGGATTTCTGTTCAATTTAGTAAATTTGAAATTTGAAATTTGAAATTTGAAATTT 1800
Qy 1801 GCAACACTAGTTTAAATTTAACTGAGTATCTCTAAGGAGTGAAGTGAAGTGAAGT 1860
Db 1801 GCAACACTAGTTTAAATTTAACTGAGTATCTCTAAGGAGTGAAGTGAAGTGAAGT 1860
Qy 1861 TTTCTGTTTAAATTTCAAGCAAACTGAGTATCTCTAAGGAGTGAAGTGAAGTGAAGT 1920
Db 1861 TTTCTGTTTAAATTTCAAGCAAACTGAGTATCTCTAAGGAGTGAAGTGAAGTGAAGT 1920
Qy 1921 AAGAAAGTTTAAATTTCAAGTATCTCTAAGGAGTGAAGTGAAGTGAAGTGAAGT 1980
Db 1921 AAGAAAGTTTAAATTTCAAGTATCTCTAAGGAGTGAAGTGAAGTGAAGTGAAGT 1980
Qy 1981 TGCATTCATATATTTTGTATAGTATAGTATAGTATAGTATAGTATAGTATAGTATAG 2040
Db 1981 TGCATTCATATATTTTGTATAGTATAGTATAGTATAGTATAGTATAGTATAGTATAG 2040
Qy 2041 AGATGCTTTGGGACCAAGAGTGTTTTGGATTTTTCAGATTTTTCAGATTTTTCAGAT 2100
Db 2041 AGATGCTTTGGGACCAAGAGTGTTTTGGATTTTTCAGATTTTTCAGATTTTTCAGAT 2100
Qy 2101 CCATACATATATGAGAGAGTGTGAAATGGGATTCAGTCTAATCATATAAATTCACCTTA 2160
Db 2101 CCATACATATATGAGAGAGTGTGAAATGGGATTCAGTCTAATCATATAAATTCACCTTA 2160
Qy 2161 TGTGATATATACCTTATCTGAATAGTCTGAAGTAAATTTTATACATATTTTAAATAA 2220
Db 2161 TGTGATATATACCTTATCTGAATAGTCTGAAGTAAATTTTATACATATTTTAAATAA 2220
Qy 2221 TTTTATGCTGAAACAGAGTGTGCGCAATTTGGACCATCAGAAAGCAGAGTGTACCTAT 2280
Db 2221 TTTTATGCTGAAACAGAGTGTGCGCAATTTGGACCATCAGAAAGCAGAGTGTACCTAT 2280
Qy 2281 TTTCAAGTCAAGTGTCTCAAAAAGTTTTCAGATGTTTGAAGTGTGATGAGTGTGATGAG 2340
Db 2281 TTTCAAGTCAAGTGTCTCAAAAAGTTTTCAGATGTTTGAAGTGTGATGAGTGTGATGAG 2340
Qy 2341 TCCGAGTACTTTGGGAGCCAAAGCAGGTGATCTTTGAGCCCAAGGAGTTTGGGCCAG 2400
Db 2341 TCCGAGTACTTTGGGAGCCAAAGCAGGTGATCTTTGAGCCCAAGGAGTTTGGGCCAG 2400

QY 2401 ACTGCACAACACAGTGAGACCTCGTTCTTCAAAATAAATAAAAAATAGCCAGGTGGT 2460
Db 2401 ACTGCACAACACAGTGAGACCTCGTTCTTCAAAATAAATAAAAAATAGCCAGGTGGT 2460
QY 2461 GGTGCACACCTGTAGTCCCGAGTACTCAGGAGGCTGAGGTAGTAGGATTTGTTGAGACTG 2520
Db 2461 GGTGCACACCTGTAGTCCCGAGTACTCAGGAGGCTGAGGTAGTAGGATTTGTTGAGACTG 2520
QY 2521 GGAGGTTGAGGCTGAACCTGAGCCAGGATCTTTGCCACACATTCAGCTTGGGCAACAGAG 2580
Db 2521 GGAGGTTGAGGCTGAACCTGAGCCAGGATCTTTGCCACACATTCAGCTTGGGCAACAGAG 2580
QY 2581 TGAGACCTGTCTCAAAATAAATAAAAAAAGTTTCAGATTTTGGAGCATTTGAGATCTTCAG 2640
Db 2581 TGAGACCTGTCTCAAAATAAATAAAAAAAGTTTCAGATTTTGGAGCATTTGAGATCTTCAG 2640
QY 2641 ATTAGGGATTTTCAACCTGTACTGACCTTTTGTAGTCATTTGACAGCATTAATCATAGGTG 2700
Db 2641 ATTAGGGATTTTCAACCTGTACTGACCTTTTGTAGTCATTTGACAGCATTAATCATAGGTG 2700
QY 2701 GACTCCAGATTAACCTGTTGCTGTATATACACATTTTGGCTCTCTATTCAAGAAATCTTAT 2760
Db 2701 GACTCCAGATTAACCTGTTGCTGTATATACACATTTTGGCTCTCTATTCAAGAAATCTTAT 2760
QY 2761 GCCCTCTTGTGGTGAATTTAATGTGCGAAGGGAACAAATAGAAATTTTGCATTTCTAGA 2820
Db 2761 GCCCTCTTGTGGTGAATTTAATGTGCGAAGGGAACAAATAGAAATTTTGCATTTCTAGA 2820
QY 2821 AAAGTCATCTGTCAAAATATGTGAGTCTGTAGATATTAGCCAAATTTAGGAAATGAC 2880
Db 2821 AAAGTCATCTGTCAAAATATGTGAGTCTGTAGATATTAGCCAAATTTAGGAAATGAC 2880
QY 2881 AAAATTTTCTTCTGCTGCTTGTAGTCTGTTTATGATATAATACCTTATTTGT 2940
Db 2881 AAAATTTTCTTCTGCTGCTTGTAGTCTGTTTATGATATAATACCTTATTTGT 2940
QY 2941 AATAAAATTAATTTTAAATTTGAGTAACAAATCGAAATTTATCAGAGAAGGGCAAGCAATA 3000
Db 2941 AATAAAATTAATTTTAAATTTGAGTAACAAATCGAAATTTATCAGAGAAGGGCAAGCAATA 3000
QY 3001 GGTAAATAACAGTATTTGATTTGAGAGGACGTTGAAATCCAGAGCATCAATGCTT 3060
Db 3001 GGTAAATAACAGTATTTGATTTGAGAGGACGTTGAAATCCAGAGCATCAATGCTT 3060
QY 3061 CTGGTGGTTTCAACCAATGAGGACGAGATGCTTAACTTTCCGAGATCTAGTTTTCAG 3120
Db 3061 CTGGTGGTTTCAACCAATGAGGACGAGATGCTTAACTTTCCGAGATCTAGTTTTCAG 3120
QY 3121 CAAAGCAGATTTAAGAAATGTAATCTTATGTTGTTTGAAGCAATAGAAATCAAT 3180
Db 3121 CAAAGCAGATTTAAGAAATGTAATCTTATGTTGTTTGAAGCAATAGAAATCAAT 3180
QY 3181 GCTGTATAAGTCTTTTAACTGTAATTTTGTGAAGCTTATCTTTTATGATATAAT 3240
Db 3181 GCTGTATAAGTCTTTTAACTGTAATTTTGTGAAGCTTATCTTTTATGATATAAT 3240
QY 3241 ATTGAAACATTTTCAATTTTATTTTAAATCAGTTTACTCAAGTGTGATTAATAC 3300
Db 3241 ATTGAAACATTTTCAATTTTATTTTAAATCAGTTTACTCAAGTGTGATTAATAC 3300
QY 3301 AAGAAATGTAACCACTGTAGGAGTATAGAAATTTTGTCAATTTGTCATTAAT 3360
Db 3301 AAGAAATGTAACCACTGTAGGAGTATAGAAATTTTGTCAATTTGTCATTAAT 3360
QY 3361 TGTAGTCACTCTCTTATGAGAGACAGAACAGTACATCTCCCAAGAAAGTTCCACAGTG 3420
Db 3361 TGTAGTCACTCTCTTATGAGAGACAGAACAGTACATCTCCCAAGAAAGTTCCACAGTG 3420
QY 3421 CTCCTTTTCCCTGAGTTTCCAGGCTGCAACCAATGATCTGCTCGTATTAATAA 3480
Db 3421 CTCCTTTTCCCTGAGTTTCCAGGCTGCAACCAATGATCTGCTCGTATTAATAA 3480
QY 3481 CTGTTCTAGATATTTGTAGCAATGTACCCTTTTCCATATTTTATTTTGTGTGTAAGGCTT 3540

Db 3481 CTGTTCTAGATATTTGTAGCAATGTACCCTTTTCCATATTTTATTTGTGTGTAAGGCTT 3540
QY 3541 CTTTGTAGTCAATTAATATTTTGTAGATTCATCTATGTTTATGTTCTATCAGTAGTTGT 3600
Db 3541 CTTTGTAGTCAATTAATATTTTGTAGATTCATCTATGTTTATGTTCTATCAGTAGTTGT 3600
QY 3601 ACATCTTACTTCTCAGCATATCACCATATAGATATATCTATAATTTTGTAAATCTAATCA 3660
Db 3601 ACATCTTACTTCTCAGCATATCACCATATAGATATATCTATAATTTTGTAAATCTAATCA 3660
QY 3661 CTGATGGATATGTAGATATTTTAACTTTTGCATTTATGAATTAAGTGGCTATTAATGAA 3720
Db 3661 CTGATGGATATGTAGATATTTTAACTTTTGCATTTATGAATTAAGTGGCTATTAATGAA 3720
QY 3721 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 3762
Db 3721 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 3762
RESULT 2
ABV20298
ID ABV20298 standard; cDNA; 1128 BP.
XX
AC ABV20298;
XX
DT 13-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 20289.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN W0200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US005171.
XX
PR 17-FEB-2000; 2000US-0183319P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JE;
XX
PS WPI; 2001-662795/76.
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS Claim 1; Page 3320-3321; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
XX

```
SQ Sequence 1128 BP; 354 A; 180 C; 216 G; 374 T; 0 U; 4 Other;
Query Match 19.5%; Score 733; DB 5; Length 1128;
Best Local Similarity 99.6%; Pred. No. 3.5e-244;
Matches 933; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1493 GTAATCTACTCAATAAAGCTGTGAAGACAGTGAACAGACCCCTTGTGAACCTGACATTT 1552
Db 58 GTAATCTACTCAATAAAGCTGTGAAGACAGTGAACAGACCCCTTGTGAACCTGACATTT 117
QY 1553 ACTTCAATTTTCTTTTCTATGACTGATATTTTGCATATAAATCTGCAGTAATAGT 1612
Db 118 ACTTCAATTTTCTTTTCTATGACTGATATTTTGCATATAAATCTGCAGTAATAGT 177
QY 1613 TCAAAAATTAATAGTTTGTGACATTTTCTGAGAAGAGAAATGAAAGTGTCAAA 1672
Db 178 TCAAAAATTAATAGTTTGTGACATTTTCTGAGAAGAGAAATGAAAGTGTCAAA 237
QY 1673 AATAAATAAGATGAATGAAGCATAATATGCTCAATTTTCAATTTTCTAGTCAAC 1732
Db 238 AATAAATAAGATGAATGAAGCATAATATGCTCAATTTTCAATTTTCTAGTCAAC 297
QY 1733 AGAATCGAAGAGTCTGTTTCAATATTAAGTAAATTTGAAAATAAATGCTTATATA 1792
Db 298 AGAATCGAAGAGTCTGTTTCAATATTAAGTAAATTTGAAAATAAATGCTTATATA 357
QY 1793 TTTTGTGTTGCAACACACATAGTAATTAATCTGCTGACTAGTATCTCTACGAGGTGGA 1852
Db 358 TTTTGTGTTGCAACACACATAGTAATTAATCTGCTGACTAGTATCTCTACGAGGTGGA 417
QY 1853 TGTGTAGTTTCTGTTTAAATTTCAAGCAATCTGGAATAATCAATCAATTAATATGCTT 1912
Db 418 TGTGTAGTTTCTGTTTAAATTTCAAGCAATCTGGAATAATCAATCAATTAATATGCTT 477
QY 1913 TCTTTCCCAAGAGTTTAAATGATGCGAGTCTCTAATTTGGAGACAAAGCCTTA 1972
Db 478 TCTTTCCCAAGAGTTTAAATGATGCGAGTCTCTAATTTGGAGACAAAGCCTTA 537
QY 1973 ATTGACATGATTCATTAATATATTTTGTATAGTTTACAGTATACAGTGTAGTATCC 2032
Db 538 ATTGACATGATTCATTAATATATTTTGTATAGTTTACAGTATACAGTGTAGTATCC 597
QY 2033 CTTAGATGAGATGCTTGGACCAAGAGTGTGTTTGGATTCAGATTTATTTTGGATTTTG 2092
Db 598 CTTAGATGAGATGCTTGGACCAAGAGTGTGTTTGGATTCAGATTTATTTTGGATTTTG 657
QY 2093 GAATATTTCCATACATATTAATGAGAGTGTGAAATGGGATTTCAAGTCTAATCAATAAAA 2152
Db 658 GAATATTTCCATACATATTAATGAGAGTGTGAAATGGGATTTCAAGTCTAATCAATAAAA 717
QY 2153 TTCACATATGTTTGATATACACCTTATCTGAATAGCCTGGAAGGTAAATTTTATACAATAT 2212
Db 718 TGCACATATGTTTGATATACACCTTATCTGAATAGCCTGGAAGGTAAATTTTATACAATAT 777
QY 2213 TTAATAATTTTATGCTGGAACAGAGTTTCCGACATGAGCACCATGAGAAAGCAGAAGT 2272
Db 778 TTAATAATTTTATGCTGGAACAGAGTTTCCGACATGAGCACCATGAGAAAGCAGAAGT 837
QY 2273 GTCATATTTCAAGTCACTGCTCAAAAGATTTTCAAGTGTAAAGCTGTGATGCAATTCAT 2332
Db 838 GTCATATTTCAAGTCACTGCTCAAAAGATTTTCAAGTGTAAAGCTGTGATGCAATTCAT 897
QY 2333 GCCAGTATCCAGTATCTTTTGGGAAGCCAGACAGGTGGATTCCTTTGAGCCCGAGGATTT 2392
Db 898 GCCAGTATCCAGTATCTTTTGGGAAGCCAGACAGGTGGATTCCTTTGAGCCCGAGGATTT 957
QY 2393 GAGGCCAGACTGCACACACACAGTGTGAGACCTGCTTTCT 2429
Db 958 GAGGCCAGACTGCACACACACAGTGTGAGACCTGCTTTCT 994

RESULT 3
ABV26130
```

```
ID ABV26130 standard; cDNA; 1128 BP.
XX AC ABV26130;
XX DT 16-SEP-2002 (first entry)
XX DE Human prostate expression marker cDNA 26121.
XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX KM pharmacogenomic marker; gene; ss.
XX OS Homo sapiens.
XX PN W0200160860-A2.
XX PD 23-AUG-2001.
XX PF 20-FEB-2001; 2001WO-US005171.
XX PR 17-FEB-2000; 2000US-0183319P.
XX PR 16-MAR-2000; 2000US-0189862P.
XX PR 25-MAY-2000; 2000US-0207454P.
XX PR 09-JUN-2000; 2000US-0211314P.
XX PR 18-JUL-2000; 2000US-0219007P.
XX PR 13-DEC-2000; 2000US-0255281P.
XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX PI Schlegel R, Endege WO, Monahan JE;
XX WPI; 2001-662795/76.
XX DR Novel isolated nucleic acid molecule associated with cancerous state of
XX PT prostate cells and correlating with presence of prostate cancer, useful
XX PT for detecting presence of prostate cancer, stage of prostate cancer.
XX PS Claim 1; Page 5285-5286; 11750pp; English.
XX CC The invention relates to an isolated nucleic acid molecule (I) comprising
XX CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX CC specification or its complement. (I) is useful for: (a) assessing whether
XX CC a patient is afflicted with prostate cancer; (b) monitoring the
XX CC progression of prostate cancer in a patient; (c) assessing the efficacy
XX CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
XX CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
XX CC (e) selecting a composition for inhibiting prostate cancer in a patient;
XX CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
XX CC determining whether prostate cancer has metastasized in a patient; (h)
XX CC assessing the aggressiveness or indolence of prostate cancer in a patient
XX CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX SQ Sequence 1128 BP; 354 A; 180 C; 216 G; 374 T; 0 U; 4 Other;
Query Match 19.5%; Score 733; DB 5; Length 1128;
Best Local Similarity 99.6%; Pred. No. 3.5e-244;
Matches 933; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1493 GTAATCTACTCAATAAAGCTGTGAAGACAGTGAACAGACCCCTTGTGAACCTGACATTT 1552
Db 58 GTAATCTACTCAATAAAGCTGTGAAGACAGTGAACAGACCCCTTGTGAACCTGACATTT 117
QY 1553 ACTTCAATTTTCTTTTCTATGACTGATATTTTGCATATAAATCTGCAGTAATAGT 1612
Db 118 ACTTCAATTTTCTTTTCTATGACTGATATTTTGCATATAAATCTGCAGTAATAGT 177
QY 1613 TCAAAAATTAATAGTTTGTGACATTTTCTGAGAAGAGAAATGAAAGTGTCAAA 1672
Db 178 TCAAAAATTAATAGTTTGTGACATTTTCTGAGAAGAGAAATGAAAGTGTCAAA 237
QY 1673 AATAAATAAGATGAATGAAGCATAATATGCTCAATTTTCAATTTTCTAGTCAAC 1732
Db 238 AATAAATAAGATGAATGAAGCATAATATGCTCAATTTTCAATTTTCTAGTCAAC 297
```


Db	598	CTTAGATGAGATGCTTTGGGACCAGAAAGTGTTTTGGATTTCAGATTATTTTGGATTTCG	657
QY	2093	GAATATTTCCATACATATATGAGAGAGTTGGAAATGGGATTTCAAAGTCTAAATCATAAAA	2152
Db	658	GAATATTTCCATACATATATGAGAGAGTTGGTAAATGGGATTTCAAAGTCTAAATCATAAGA	717
QY	2153	TTCACTTATGTTTGAATATACACCTTATCTGAATAGCCTGAAAGTAAATTTTATACAATATT	2212
Db	718	TGCACTTATGTTTGAATATACACCTTATCTGAATAGCCTGAAAGTAAATTTTATACAATATT	777
QY	2213	TTAAATAATTTTATGCTGCTGAAACACAGAGTTTTCGCGACATTCGACCATCAGAAACGACAGAAGT	2272
Db	778	TTAAATAATTTTATGCTGCTGAAACACAGAGTTTTCGCGACATTCGACCATCAGAAACGACAGAAGT	837
QY	2273	GTCACTATTTCAAAGTCAAGTCTCAAAAAGTTTCAGATGTTTAAAGCTGGTGATGCAGTTTCAT	2332
Db	838	GTCACTATTTCAAAGTCAAGTCTCAAAAAGTTTCAGATGTTTAAAGCTGGTGATGCAGTTTCAT	897
QY	2333	GCCAGTATCCGAGTACTTTTGGGAAGCCCAAGACAGAGTGGATCTCTTGAGCCCGAGGAGTTT	2392
Db	898	GCCAGTATCCGAGTACTTTTGGGAAGCCCAAGACAGAGTGGATCTCTTGAGCCCGAGGAGTTT	957
QY	2393	GAGGCCAGAGCTGCACAACACAGTGCAGACCTCGTTTCT	2429
Db	958	GAGGCCAGAGCTGCACAACACAGTGCAGACCTCGTTTCT	994
RESULT 5			
ABV20617	ID ABV20617 standard; cDNA; 1128 BP.		
XX	XX		
AC	ABV20617;		
XX	XX		
DT	13-SEP-2002 (first entry)'		
XX	XX		
DE	Human prostate expression marker cDNA 20608.		
XX	XX		
KW	Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;		
XX	XX		
OS	Homo sapiens.		
XX	XX		
FN	WO200160860-A2.		
XX	XX		
PD	23-AUG-2001.		
XX	XX		
PF	20-FEB-2001; 2001WO-US005171.		
XX	XX		
PR	17-FEB-2000; 2000US-0183319P.		
PR	16-MAR-2000; 2000US-0189862P.		
PR	25-MAY-2000; 2000US-0207454P.		
PR	09-JUN-2000; 2000US-0211314P.		
PR	18-JUL-2000; 2000US-0219007P.		
PR	13-DEC-2000; 2000US-0255281P.		
XX	(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.		
PA	XX		
XX	Schlegel R, Endege WO, Monahan JE;		
PI	XX		
XX	WPI; 2001-662795/76.		
DR	XX		
XX	Novel isolated nucleic acid molecule associated with cancerous state of		
PT	prostate cells and correlating with presence of prostate cancer, useful		
PT	for detecting presence of prostate cancer, stage of prostate cancer.		
XX	XX		
PS	Claim 1; Page 3379; 11750pp; English.		
XX	XX		
CC	The invention relates to an isolated nucleic acid molecule (I) comprising		
CC	a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the		
CC	specification or its complement. (I) is useful for: (a) assessing whether		
CC	a patient is afflicted with prostate cancer; (b) monitoring the		
CC	progression of prostate cancer in a patient; (c) assessing the efficacy		
CC	of a test compound to inhibit prostate cancer in a patient; (d) assessing		

QY 2393 GAGCCGACGCTGCACACACACAGTGAGACCTCGTTTCT 2429
 DB 958 GAGCCGACGCTGCACACACACAGTGAGACCTCGTTTCT 994

RESULT 6

ID ABV20806 standard; cDNA; 1128 BP.

AC ABV20806;

XX 13-SEP-2002 (first entry)

XX Human prostate expression marker cDNA 20797.

XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

XX pharmacogenomic marker; gene; ss.

XX Homo sapiens.

XX WO200160860-A2.

XX 23-AUG-2001.

XX 20-FEB-2001; 2001WO-US005171.

XX 17-FEB-2000; 2000US-0183319P.

XX 16-MAR-2000; 2000US-0189862P.

XX 25-MAY-2000; 2000US-0207454P.

XX 09-JUN-2000; 2000US-0211314P.

XX 18-JUL-2000; 2000US-0213007P.

XX 13-DEC-2000; 2000US-0255281P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Schlegel R, Endege WO, Monahan JE;

XX WPI; 2001-662795/76.

XX Novel isolated nucleic acid molecule associated with cancerous state of

XX prostate cells and correlating with presence of prostate cancer, useful

XX for detecting presence of prostate cancer, stage of prostate cancer.

XX Claim 1; Page 3412-3413; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising

XX a nucleotide sequence given in Tables 1-9 (ABV0010-ABV62213) of the

XX specification or its complement. (I) is useful for: (a) assessing whether

XX a patient is afflicted with prostate cancer; (b) monitoring the efficacy

XX of a test compound to inhibit prostate cancer in a patient; (c) assessing the efficacy

XX of a test compound to inhibit prostate cancer in a patient; (d) assessing

XX the efficacy of a therapy for inhibiting prostate cancer in a patient;

XX (e) selecting a composition for inhibiting prostate cancer in a patient;

XX (f) assessing the prostate cell carcinogenic potential of a compound; (g)

XX determining whether prostate cancer has metastasized in a patient; (h)

XX assessing the aggressiveness or indolence of prostate cancer in a patient

XX ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker

XX

XX Query Match 19.5%; Score 733; DB 5; Length 1128;

XX Best Local Similarity 99.6%; Pred. No. 3.5e-244;

XX Matches 933; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1493 GTAATCTACTCAATAAAGCTGTATAGACACAGTGACAGACCTTTGCTAACTGACATTT 1552

DB 58 GTAATCTACTCAATAAAGCTGTATAGACACAGTGACAGACCTTTGCTAACTGACATTT 117

QY 1553 ACTTCAATTTTCTTTTCTATGTACTGTATATTTTGCATATAAATTTGAGTAATAGT 1612

DB 118 ACTTCAATTTTCTTTTCTATGTACTGTATATTTTGCATATAAATTTGAGTAATAGT 177

QY 1613 TCAAAAATTAATAGTATTTTTCATGATTTTCTCGAAGAGAGAAATGAAAGTGTCACAA 1672

DB 178 TCAAAAATTAATAGTATTTTTCATGATTTTCTCGAAGAGAGAAATGAAAGTGTCACAA 237
 QY 1673 AATAAAAAAGATGAAATGAAGCATATATAATTTGTCATTTTTCATTTTCTAGTCAAC 1732
 DB 238 AATAAAAAAGATGAAATGAAGCATATATAATTTGTCATTTTTCATTTTCTAGTCAAC 297
 QY 1733 AGAGATCGAAGATCTGTTCAAAATATAGTAAAAATGAAAAATGAAATCTGTGCTTATA 1792
 DB 298 AGAGATCGAAGATCTGTTCAAAATATAGTAAAAATGAAAAATGAAATCTGTGCTTATA 357
 QY 1793 TTTTGTTCGCAACACTAGTTTAAATTTAACCTGTGCTAGTTTATCTTACCGAAGGTGA 1852
 DB 358 TTTTGTTCGCAACACTAGTTTAAATTTAACCTGTGCTAGTTTATCTTACCGAAGGTGA 417
 QY 1853 TGTGTAGTTTCTGTTTAAATTTCAAGCAAACTCGAAAAATTAATCCATCTAATATGCTT 1912
 DB 418 TGTGTAGTTTCTGTTTAAATTTCAAGCAAACTCGAAAAATTAATCCATCTAATATGCTT 477
 QY 1913 TCTTTCCCAAGAGTTTAAATTTAATGATGCGCAGCTTCTTAATTTGGAGACAAAAGCCTTA 1972
 DB 478 TCTTTCCCAAGAGTTTAAATTTAATGATGCGCAGCTTCTTAATTTGGAGACAAAAGCCTTA 537
 QY 1973 ATTGACATGCAATTCATTATATATTTTGTATAGTTTACAGTATACGAGTTGATGCTC 2032
 DB 538 ATTGACATGCAATTCATTATATATTTTGTATAGTTTACAGTATACGAGTTGATGCTC 597
 QY 2033 CTTAGATGAGATGCTTGGACAGAGGTTTGGATTTTGGATTTTGGATTTTGGATTTTGG 2092
 DB 598 CTTAGATGAGATGCTTGGACAGAGGTTTGGATTTTGGATTTTGGATTTTGGATTTTGG 657
 QY 2093 GAATATTTCCATACATATATTAATGAGAGAGTTGGAAAAATGGATTTCAAGTCTAATCAAAA 2152
 DB 658 GAATATTTCCATACATATTAATGAGAGAGTTGGTAAATGGATTTCAAGTCTAATCAAAA 717
 QY 2153 TTCACTTATGTTGATATACACCTTATCTGAATAGCTCAAGTAAATTTTATACATATT 2212
 DB 718 TGCACTTATGTTGATATACACCTTATCTGAATAGCTCAAGTAAATTTTATACATATT 777
 QY 2213 TTAATAATTTTATGCTGAAACAGAGTTTGGCCACATTTGGACCATCAGAAAGCAGAGT 2272
 DB 778 TTAATAATTTTATGCTGAAACAGAGTTTGGCCACATTTGGACCATCAGAAAGCAGAGT 837
 QY 2273 GTCATATTTCAAGTCAGTGCTCAAAAAGTTTCAGATGTTAAGCTGGTGTGATGCTTCAAT 2332
 DB 838 GTCATATTTCAAGTCAGTGCTCAAAAAGTTTCAGATGTTAAGCTGGTGTGATGCTTCAAT 897
 QY 2333 GCCAGTGATCCGAGTACTTTTGGAGCCCAAGACAGTGATCTCTTGGAGCCAGGAGTTT 2392
 DB 898 GCCAGTGATCCGAGTACTTTTGGAGCCCAAGACAGTGATCTCTTGGAGCCAGGAGTTT 957
 QY 2393 GAGGCCAGACTGCACAAACACAGTGAGACCTCGTTTCT 2429
 DB 958 GAGGCCAGACTGCACAAACACAGTGAGACCTCGTTTCT 994

RESULT 7

ABV20439

ID ABV20439 standard; cDNA; 1128 BP.

XX AC ABV20439;

XX 13-SEP-2002 (first entry)

XX Human prostate expression marker cDNA 20430.

XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

XX pharmacogenomic marker; gene; ss.

XX Homo sapiens.

XX WO200160860-A2.

XX

PD 23-AUG-2001.
 XX 20-FEB-2001; 2001WO-US005171.
 XX 17-FEB-2000; 2000US-0183319P.
 XX 16-MAR-2000; 2000US-0189862P.
 XX 25-MAY-2000; 2000US-0207454P.
 XX 09-JUN-2000; 2000US-0211314P.
 XX 18-JUL-2000; 2000US-0219007P.
 XX 13-DEC-2000; 2000US-0255281P.
 XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX Schlegel R, Endege WO, Monahan JE;
 XX WPI; 2001-662795/76.
 PT Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer.
 XX
 PS Claim 1; Page 3347; 11750pp; English.
 CC The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for: (a) assessing whether
 CC a patient is afflicted with prostate cancer; (b) monitoring the
 CC progression of prostate cancer in a patient; (c) assessing the efficacy
 CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
 CC determining whether prostate cancer has metastasized in a patient; (h)
 CC assessing the aggressiveness or indolence of prostate cancer in a patient
 CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
 XX
 SQ Sequence 1128 BP; 354 A; 180 C; 216 G; 374 T; 0 U; 4 Other;
 Query Match 19.5%; Score 733; DB 5; Length 1128;
 Best Local Similarity 99.6%; Pred. No. 3.5e-244;
 Matches 933; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1493 GTAATCTACTCAATAAAGCTGTGAAGACAGTACGACACCTTTGTAACCTGACATTT 1552
 DB 58 GTAATCTACTCAATAAAGCTGTGAAGACAGTACGACACCTTTGTAACCTGACATTT 117
 QY 1553 ACTTCAATTTCTTTTCTATGCTGATATTTTGCATATAAATCTGCAAGTAAGT 1612
 DB 118 ACTTCAATTTCTTTTCTATGCTGATATTTTGCATATAAATCTGCAAGTAAGT 177
 QY 1613 TCATAAATTAATAGTTTGTGATGCTTTTCTGAGAAGAGAAATGAAAGTGTCAAA 1672
 DB 178 TCATAAATTAATAGTTTGTGATGCTTTTCTGAGAAGAGAAATGAAAGTGTCAAA 237
 QY 1673 AATAAATAAGATGAATGAGCATATATATTTGTCATTTTTTCAATTTTCTAGTCAAC 1732
 DB 238 AATAAATAAGATGAATGAGCATATATATTTGTCATTTTTTCAATTTTCTAGTCAAC 297
 QY 1733 AGAGAATCGAAGGATTCGTGTTCAATATAGTAAATAATGAAATAAATCTGCTTATA 1792
 DB 298 AGAGAATCGAAGGATTCGTGTTCAATATAGTAAATAATGAAATAAATCTGCTTATA 357
 QY 1793 TTTGTTTGCACACATAGTTAATTAACCTGTGACTAGTATCTCTACCGAAGGTGGA 1852
 DB 358 TTTGTTTGCACACATAGTTAATTAACCTGTGACTAGTATCTCTACCGAAGGTGGA 417
 QY 1853 TGTGTAGTTCTGTTTAAATAATCAAGCAACTGGAATAATCAATCTAATATATGCTT 1912
 DB 418 TGTGTAGTTCTGTTTAAATAATCAAGCAACTGGAATAATCAATCTAATATATGCTT 477
 QY 1913 TCTTTTCCCAAGAGTTTTTAAATGATATGCAAGCTTCTTAATTTGGAGACAAAGCCTTA 1972
 DB 478 TCTTTTCCCAAGAGTTTTTAAATGATATGCAAGCTTCTTAATTTGGAGACAAAGCCTTA 537

QY 1973 ATTGCAATGCAATTCATTATATATTTTTTTGTATAGTTTACAGTATACGAGTTGAGTATCC 2032
 DB 538 ATTGCAATGCAATTCATTATATATTTTTTTGTATAGTTTACAGTATACGAGTTGAGTATCC 597
 QY 2033 CTTAGATGAGATGCTTGGGACGAGAGTGTGTTTGGATTTCAGATTATTTTGGATTG 2092
 DB 598 CTTAGATGAGATGCTTGGGACGAGAGTGTGTTTGGATTTCAGATTATTTTGGATTG 657
 QY 2093 GAATATTTCCATACATATAATGAGAGAGTTGGAATGGATTCAAGTCTAATCATATAAA 2152
 DB 658 GAATATTTCCATACATATAATGAGAGAGTTGGAATGGATTCAAGTCTAATCATATAAA 717
 QY 2153 TTCACATATGTTGATATACACCTTATCTGAATAGCCTGAGGTAAATTTTATACAAATTT 2212
 DB 718 TGCACATATGTTGATATACACCTTATCTGAATAGCCTGAGGTAAATTTTATACAAATTT 777
 QY 2213 TTAATAATTTTTATGCTGAAACAGAGTTTGGCAGCATTTGGACCATCAGAAAGCAGAGT 2272
 DB 778 TTAATAATTTTTATGCTGAAACAGAGTTTGGCAGCATTTGGACCATCAGAAAGCAGAGT 837
 QY 2273 GTCACATTTTCAAGTCAGTGTCTCAAAAAGTTTTCAGATGTTAAGCTGGTATGCAATTCAT 2332
 DB 838 GTCACATTTTCAAGTCAGTGTCTCAAAAAGTTTTCAGATGTTAAGCTGGTATGCAATTCAT 897
 QY 2333 GCCAGTATCCGAGTACTTTGGGAGGCCAAGACAGAGTGTCTCTTGAGCCCGAGGATTT 2392
 DB 898 GCCAGTATCCGAGTACTTTGGGAGGCCAAGACAGAGTGTCTCTTGAGCCCGAGGATTT 957
 QY 2393 GAGGCCAGACTCTCACAACACACAGTGTGAGCTCGTTTCT 2429
 DB 958 GAGGCCAGACTCTCACAACACACAGTGTGAGCTCGTTTCT 994
 RESULT 8
 ABV26275
 ID ABV26275 standard; cDNA; 1128 BP.
 XX AC ABV26275;
 XX AC ABV26275;
 DT 16-SEP-2002 (first entry)
 XX Human prostate expression marker cDNA 26266.
 DE Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.
 OS Homo sapiens.
 XX WO200160860-A2.
 XX 23-AUG-2001.
 XX 20-FEB-2001; 2001WO-US005171.
 XX 17-FEB-2000; 2000US-0183319P.
 XX 16-MAR-2000; 2000US-0189862P.
 XX 25-MAY-2000; 2000US-0207454P.
 XX 09-JUN-2000; 2000US-0211314P.
 XX 18-JUL-2000; 2000US-0219007P.
 XX 13-DEC-2000; 2000US-0255281P.
 XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX Schlegel R, Endege WO, Monahan JE;
 XX WPI; 2001-662795/76.
 XX Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer.
 XX
 PS Claim 1; Page 5313; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 1128 BP; 354 A; 180 C; 216 G; 374 T; 0 U; 4 Other;

Query Match 19.5%; Score 733; DB 5; Length 1128;
Best Local Similarity 99.6%; Pred. No. 3.5e-244;
Matches 933; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1493 GTAATCTACTCAATAAAGCTGTAGACAGCTGACAGAGCCCTTGTCAACTGACATTT 1552
Db |||||
QY 58 GTAATCTACTCAATAAAGCTGTAGACAGCTGACAGAGCCCTTGTCAACTGACATTT 117
Db |||||
QY 1553 ACTTCAATTTTCTTTTCTATGTACTGATATTTTGCATATATAAATTCGAGTAATAGT 1612
Db |||||
QY 1613 TCAAAATTAATAGTTTTCGATTTGGCTTTCTGAGAGAGAAATGAAAGTGCACAA 1672
Db |||||
QY 178 TCAAAATTAATAGTTTTCGATTTGGCTTTCTGAGAGAGAAATGAAAGTGCACAA 237
Db |||||
QY 1673 AATAAATAAGATGAAATGAACATATATAATGTCAATTTTCAATTTTCTAGTCAAC 1732
Db |||||
QY 238 AATAAATAAGATGAAATGAACATATATAATGTCAATTTTCAATTTTCTAGTCAAC 297
Db |||||
QY 1733 AGAGATCGAAGATTCGTTCAATATATAGTAAATTAATGAAATTAATGCTTATATA 1792
Db |||||
QY 298 AGAGATCGAAGATTCGTTCAATATATAGTAAATTAATGAAATTAATGCTTATATA 357
Db |||||
QY 1793 TTTTGTGTCACACACTAGTTTAAATTAACCTGTGACTAGTTATCTCTACCGAGAGTGA 1852
Db |||||
QY 358 TTTTGTGTCACACACTAGTTTAAATTAACCTGTGACTAGTTATCTCTACCGAGAGTGA 417
Db |||||
QY 1853 TGTGATGTTCTGGTTTAAATTAACCTGTGACTAGTTATCTCTACCGAGAGTGA 1912
Db |||||
QY 418 TGTGATGTTCTGGTTTAAATTAACCTGTGACTAGTTATCTCTACCGAGAGTGA 477
Db |||||
QY 1913 TCTTTCCCAAGAGTTTAAATTAACCTGTGACTAGTTATCTCTACCGAGAGTGA 1972
Db |||||
QY 478 TCTTTCCCAAGAGTTTAAATTAACCTGTGACTAGTTATCTCTACCGAGAGTGA 537
Db |||||
QY 1973 AATGCAATGCAATTCATTTATATATATTTTGTATAGTTTACAGTATGAGTATCC 2032
Db |||||
QY 538 AATGCAATGCAATTCATTTATATATATTTTGTATAGTTTACAGTATGAGTATCC 597
Db |||||
QY 2033 CTTAGATGAGTCTGGGACAGAGTGTGTTGGATTCAGATTTATTTTGTGATTTG 2092
Db |||||
QY 598 CTTAGATGAGTCTGGGACAGAGTGTGTTGGATTCAGATTTATTTTGTGATTTG 657
Db |||||
QY 2093 GAATATTTCCATATATATATGAGAGTGTGAAATGCGATTCAGATTCATATATAA 2152
Db |||||
QY 658 GAATATTTCCATATATATGAGAGTGTGAAATGCGATTCAGATTCATATATAA 717
Db |||||
QY 2153 TTTCACTTATGTTGATATACACTTATCTGATAGCTGAGGTAATTTTATACATTT 2212
Db |||||
QY 718 TGCACCTATGTTGATATACACTTATCTGATAGCTGAGGTAATTTTATACATTT 777
Db |||||
QY 2213 TTAATAATTTTATGCTGAAACAGAGTTTGGCAGATTCAGATTCAGATTCAGATTC 2272
Db |||||
QY 778 TTAATAATTTTATGCTGAAACAGAGTTTGGCAGATTCAGATTCAGATTCAGATTC 837
Db |||||
QY 2273 GTCACATTTTCAAGTCAGTCTCAAAAAGTTTCAGATTTTCAAGTGTGATGCAATTCAT 2332
Db |||||

Db 838 GTCACATTTTCAAGTCAGTCTCAAAAAGTTTCAGATTTTCAAGTGTGATGCAATTCAT 897
QY 2333 GCCAGTGTATCCAGTACTTTTGGAGAGCCAAAGACAGTGTGATCTCTTTGAGCCAGAGTTT 2392
Db |||||
QY 898 GCCAGTGTATCCAGTACTTTTGGAGAGCCAAAGACAGTGTGATCTCTTTGAGCCAGAGTTT 957
Db |||||
QY 2393 GAGGCCAGACTGACAAACACACAGTGTGATCTCTTTTCT 2429
Db |||||
QY 958 GAGGCCAGACTGACAAACACACAGTGTGATCTCTTTTCT 994
Db |||||

RESULT 9
ABV26653
ID ABV26653 standard; cDNA; 1128 BP.
XX AC ABV26653;
XX XX
DT 16-SBP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 26644.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US005171.
XX
PR 17-FEB-2000; 2000US-0183319P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JE;
XX
XX WPI; 2001-662795/76.
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
prostate cells and correlating with presence of prostate cancer, useful
for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS Claim 1; Page 5383; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
specification or its complement. (I) is useful for: (a) assessing whether
a patient is afflicted with prostate cancer; (b) monitoring the
progression of prostate cancer in a patient; (c) assessing the efficacy
of a test compound to inhibit prostate cancer in a patient; (d) assessing
the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 1128 BP; 354 A; 180 C; 216 G; 374 T; 0 U; 4 Other;

Query Match 19.5%; Score 733; DB 5; Length 1128;
Best Local Similarity 99.6%; Pred. No. 3.5e-244;
Matches 933; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1493 GTAATCTACTCAATAAAGCTGTAGACAGCTGACAGAGCCCTTGTCAACTGACATTT 1552
Db |||||

Db 58 GTAATCTACTCAATAAACTGTGAAGACCACTGACACAGCCCTTGCTAACTGACATTT 117
Qy 1553 ACTTCAATTTTCTTTTCTATGCTACTGATATTTTGCATATAAACTTGCAGTATAGT 1612
Db 118 ACTTCAATTTTCTTTTCTATGCTACTGATATTTTGCATATAAACTTGCAGTATAGT 177
Qy 1613 TCAAAAATTAATAGTTTTCACATTCGCTTTTCTGAGAAGAGAAATTTGAAGTGTGCACAA 1672
Db 178 TCAAAAATTAATAGTTTTCACATTCGCTTTTCTGAGAAGAGAAATTTGAAGTGTGCACAA 237
Qy 1673 AATAAARAGATGAATGAAGCATATATAATTTGTCATATTTTCAATTTTCTAGTCAAC 1732
Db 238 AATAAARAGATGAATGAAGCATATATAATTTGTCATATTTTCAATTTTCTAGTCAAC 297
Qy 1733 AGAATTCGAAGGATTCGTTTCAAAATTAAGTAAATAATGAAAATTAACCTTGCTTATA 1792
Db 298 AGAATTCGAAGGATTCGTTTCAAAATTAAGTAAATAATGAAAATTAACCTTGCTTATA 357
Qy 1793 TTTTGTGTTGCAACACACTAGTAAATTTAACTGTGACTAGTATCTCTACCGAAGTGA 1852
Db 358 TTTTGTGTTGCAACACACTAGTAAATTTAACTGTGACTAGTATCTCTACCGAAGTGA 417
Qy 1853 TGTGTAGTTCTGCTTTTAAATTTCAAGCAAACTGGAATAATCCATCTAAATTAATGCTT 1912
Db 418 TGTGTAGTTCTGCTTTTAAATTTCAAGCAAACTGGAATAATCCATCTAAATTAATGCTT 477
Qy 1913 TCTTTCCCAAGAGTTTAAATGATGATGCGAGCTTCTTAATTTGGAGACAAAGCTTA 1972
Db 478 TCTTTCCCAAGAGTTTAAATGATGATGCGAGCTTCTTAATTTGGAGACAAAGCTTA 537
Qy 1973 ATTGACATGCAATTCATATATATTTTGTATAGTTTACAGTATACGAGTTCAGTATCC 2032
Db 538 ATTGACATGCAATTCATATATATTTTGTATAGTTTACAGTATACGAGTTCAGTATCC 597
Qy 2033 CTTAGATGAGATGCTTGGGACCAAGAGTTTTCGATTTTTCAGATTTATTTTTCGATTTT 2092
Db 598 CTTAGATGAGATGCTTGGGACCAAGAGTTTTCGATTTTTCAGATTTATTTTTCGATTTT 657
Qy 2093 GAATATTTCCATACATATATGAGAGAGTTGGAAATGGGATTCAGTCTAAATCAATAAA 2152
Db 658 GAATATTTCCATACATATATGAGAGAGTTGGTAAATGGGATTCAGTCTAAATCAATAAA 717
Qy 2153 TTCCTATGTTTGTATATACACCTTATCTGAATAGCCTGAGGTAATTTTATCAATATT 2212
Db 718 TGCCTTATGTTTGTATATACACCTTATCTGAATAGCCTGAGGTAATTTTATCAATATT 777
Qy 2213 TTAATATATTTTATGCTGAAACAGAGTTTGGCAGATTCGACCATCAGAAAGCAAGT 2272
Db 778 TTAATATATTTTATGCTGAAACAGAGTTTGGCAGATTCGACCATCAGAAAGCAAGT 837
Qy 2273 GTCACTATTTCAAGTCAAGTGTCTCAAAAAGTTTCAGATGTTTAAAGTGTGATGAGTTTCAT 2332
Db 838 GTCACTATTTCAAGTCAAGTGTCTCAAAAAGTTTCAGATGTTTAAAGTGTGATGAGTTTCAT 897
Qy 2333 GCCAGTATCCAGTACTTTTGGGAAGCCAGACAGAGTGGATCTCTTGGAGCCAGGAGTTT 2392
Db 898 GCCAGTATCCAGTACTTTTGGGAAGCCAGACAGAGTGGATCTCTTGGAGCCAGGAGTTT 957
Qy 2393 GAGGCCAGACTGCAACACACAGTGCAGACCTCGTTTCT 2429
Db 958 GAGGCCAGACTGCAACACACAGTGCAGACCTCGTTTCT 994

RESULT 10

ADC30760

ID ADC30760 standard; cDNA; 667 BP.

AC ADC30760;

XX 18-DEC-2003 (first entry)

DT 18-DEC-2003 (first entry)

XX Human novel cDNA sequence, SEQ ID NO:842.

DE Human novel cDNA sequence, SEQ ID NO:842.

XX

KW Human; diagnostic; drug screening; forensics; gene mapping;
KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;
KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
KW ulcers; osteoporosis; autoimmune disease; cancer;
KW molecular weight marker; food supplement; antiparkinsonian; nootropic;
KW neuroprotective; anti-anaemic; anticoagulant; thrombolytic; vulnerary;
KW antitumor; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
KW gene therapy; chromosome 1; gene; ss.
XX Homo sapiens.
XX WO2003029271-A2.
XX 10-APR-2003.
XX 24-SEP-2002; 2002WO-US030474.
XX 24-SEP-2001; 2001US-0324631P.
XX (HYSE-) HYSEQ INC.
XX Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
PI Haley-Vicente D, Drmanac RT;
XX WPI; 2003-371981/35.
DR F-PSDB; ADC31731.

New polynucleotide and polypeptide useful for diagnosing, preventing or treating conditions such as neurodegenerative diseases, anemias, platelet disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or cancer.

Claim 1; SEQ ID NO 842; 1185pp; English.

The invention relates to 971 novel human cDNA sequences (ADC29919-ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The invention also relates to nucleic acid sequences over 99% identical with the novel human cDNAs. The invention additionally encompasses expression vectors and host cells comprising a nucleic acid of the invention; the recombinant production of a polypeptide of the invention; an antibody against a polypeptide of the invention; a method of detecting polynucleotides or polypeptides of the invention; and methods of identifying a compound which binds to a polypeptide of the invention. The invention further discloses methods of preventing, treating or ameliorating a medical condition; kits comprising polynucleotide probes and/or monoclonal antibodies for carrying out the methods of the invention; methods for the identification of compounds that modulate the expression or activity of the polynucleotide and/or polypeptide; and 787 config sequences corresponding to the cDNA sequences of the invention (ADC31861-ADC32627) and the polypeptides encoded by the configs (ADC32628-ADC33394). The nucleic acids and polypeptides of the invention are useful in diagnostics, drug screening, forensics, gene mapping, in the identification of mutations responsible for genetic disorders or other traits, for assessing biodiversity, and in producing many other types of data and products dependent on DNA and amino acid sequences. They are also used for treating diseases such as Parkinson's disease, Alzheimer's disease and other neurodegenerative diseases, anaemia, platelet disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or cancer. The nucleic acids may also be used as hybridization probes or primers, and in the recombinant production of a protein. The polypeptides are also useful in generating antibodies, as molecular weight markers, and as food supplements. The present sequence represents a specifically claimed human cDNA sequence of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 667 BP; 216 A; 147 C; 154 G; 150 T; 0 U; 0 Other;

Query Match 13.5%; Score 506; DB 9; Length 667;

Best Local Similarity 100.0%; Pred. No. 1.4e-165;

Matches 506; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGAAACCGAGTCACTGTGAAAGAGTGGAAATTTATCTCTGCGAAATCACTCAGTTGCTG 60
 DB |||||
 QY 162 GAGAAACCGAGTCACTGTGAAAGAGTGGAAATTTATCTCTGCGAAATCACTCAGTTGCTG 221
 DB |||||
 QY 61 GGAGAGATCAAAAGACCCAGAGAAACCCAGATGAGGAAAGAAAGAAAGCGCAGAA 120
 DB |||||
 QY 222 GGAGAGATCAAAAGACCCAGAGAAACCCAGATGAGGAAAGAAAGAAAGCGCAGAA 281
 DB |||||
 QY 121 ATGACTACATTTGAAAGAAACTTCAAGATCAAGATAGAAAGCCAGAAAGTTTCATCC 180
 DB |||||
 QY 282 ATGACTACATTTGAAAGAAACTTCAAGATCAAGATAGAAAGCCAGAAAGTTTCATCC 341
 DB |||||
 QY 181 ACTTCTAATCAGAAACGAGATGGAGTGTCTGAGAAAGTGTGCTACACTGTCAAT 240
 DB |||||
 QY 342 ACTTCTAATCAGAAACGAGATGGAGTGTCTGAGAAAGTGTGCTACACTGTCAAT 401
 DB |||||
 QY 241 AATCATATCCCATCAGAGATCCCTGCTGAGTCCATGATGATGCTATGAGAAAT 300
 DB |||||
 QY 402 AATCATATCCCATCAGAGATCCCTGCTGAGTCCATGATGATGCTATGAGAAAT 461
 DB |||||
 QY 301 GATCCCTCACAAGAAAGTGAAGATTTAGAGAAAGTGAAGAAAGTGAAGAAATGACCTT 360
 DB |||||
 QY 462 GATCCCTCACAAGAAAGTGAAGATTTAGAGAAAGTGAAGAAAGTGAAGAAATGACCTT 521
 DB |||||
 QY 361 CTTAGGACTTCTGTTAGTAGGCTTGTCTGCTGACCCATGAGCATGATGATGAGTTGTG 420
 DB |||||
 QY 522 CTTAGGACTTCTGTTAGTAGGCTTGTCTGCTGACCCATGAGCATGATGATGAGTTGTG 581
 DB |||||
 QY 421 TTTCACACATAAAATCCTCAAGCTCTTTATCACCTTCCAGCAATGAAGACATGCGAGAA 480
 DB |||||
 QY 582 TTTCACACATAAAATCCTCAAGCTCTTTATCACCTTCCAGCAATGAAGACATGCGAGAA 641
 DB |||||
 QY 481 TAGCGAGCTCTGGGAAAGTTGTTAC 506
 DB |||||
 QY 642 TAGCGAGCTCTGGGAAAGTTGTTAC 567
 DB |||||

RESULT 11

ABV10439

ID ABV10439 standard; cDNA; 397 BP.

XX AC ABV10439;

XX DT 13-SEP-2002 (first entry)

XX DE Human prostate expression marker cDNA 10430.

XX DE Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

XX KW pharmacogenomic marker; gene; ss.

XX OS Homo sapiens.

XX PN WO200160860-A2.

XX PD 23-AUG-2001.

XX PF 20-FEB-2001; 2001WO-US005171.

XX PR 17-FEB-2000; 2000US-0183319P.

XX PR 16-MAR-2000; 2000US-0189862P.

XX PR 25-MAY-2000; 2000US-0207454P.

XX PR 09-JUN-2000; 2000US-0211314P.

XX PR 18-JUL-2000; 2000US-0219007P.

XX PR 13-DEC-2000; 2000US-0255281P.

XX PA (WILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX PI Schlegel R, Endege WO, Monahan JE;

XX DR WPI; 2001-662795/76.

XX PT Novel isolated nucleic acid molecule associated with cancerous state of

PT prostate cells and correlating with presence of prostate cancer, useful
 for detecting presence of prostate cancer, stage of prostate cancer.
 XX Claim 1; Page 1680; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising
 a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 specification or its complement. (I) is useful for: (a) assessing whether
 a patient is afflicted with prostate cancer; (b) monitoring the
 progression of prostate cancer in a patient; (c) assessing the efficacy
 of a test compound to inhibit prostate cancer in a patient; (d) assessing
 the efficacy of a therapy for inhibiting prostate cancer in a patient;
 (e) selecting a composition for inhibiting prostate cancer in a patient;
 (f) assessing the prostate cell carcinogenic potential of a compound; (g)
 determining whether prostate cancer has metastasized in a patient; (h)
 assessing the aggressiveness or indolence of prostate cancer in a patient
 CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker

SQ Sequence 397 BP; 142 A; 53 C; 64 G; 138 T; 0 U; 0 Other;

Query Match 9.1%; Score 344; DB 5; Length 397;

Best Local Similarity 99.7%; Pred. No. 1.8e-109;

Matches 394; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1575 GTACTCGATATTTTGCATATAAACTTGCAGTAATAGTTCAAAAATTAATAGTTTTCAC 1634

DB 3 GTACTCGATATTTTGCATATAAACTTGCAGTAATAGTTCAAAAATTAATAGTTTTCAC 62

QY 1635 ATTGGCTTTTCTGAGAAGAGAAATGAAAGTGTCAAAAATTAATAAGATGAAG 1694

DB 63 ATTGGCTTTTCTGAGAAGAGAAATGAAAGTGTCAAAAATTAATAAGATGAAG 122

QY 1695 CATATATAATTTCTCAATTTTTCATTTTCTAGTCAACAGAGATCGAGGATTCGTTC 1754

DB 123 CATATATAATTTCTCAATTTTTCATTTTCTAGTCAACAGAGATCGAGGATTCGTTC 182

QY 1755 AAATATTAGTAAATAATGAAATAAACTTGTCTATATTTTGTTCACACACTAGTT 1814

DB 183 AAATATTAGTAAATAATGAAATAAACTTGTCTATATTTTGTTCACACACTAGTT 242

QY 1815 AATTTAACTGTGACTAGTATCTCTACCAAGGTGGATGCTAGTTCTGTTTTTAA 1874

DB 243 AATTTAACTGTGACTAGTATCTCTACCAAGGTGGATGCTAGTTCTGTTTTTAA 302

QY 1875 TTCAAGCAAACTGGAAAAATAATCCATCTAATATGCTTTCTTCCCAAGAGTTTTTAA 1934

DB 303 TTCAAGCAAACTGGAAAAATAATCCATCTAATATGCTTTCTTCCCAAGAGTTTTTAA 362

QY 1935 TGATATGCCAGCTTCTCTAATTTTGGAGACAAAAGCC 1969

DB 363 TGATATGCCAGCTTCTCTAATTTTGGAGACAAAAGCC 397

RESULT 12

ABV01270

ID ABV01270 standard; cDNA; 433 BP.

XX AC ABV01270;

XX DT 13-SEP-2002 (first entry)

XX DE Human prostate expression marker cDNA 1261.

XX DE Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

XX KW pharmacogenomic marker; gene; ss.

XX OS Homo sapiens.

XX PN WO200160860-A2.

XX PD 23-AUG-2001.

XX PF 20-FEB-2001; 2001WO-US005171.

QY 1875 TTCAAGCAAACTGGAAATAATCAATCAATTAATGCTTTCTTCCCAAGAGTTTTTAA 1934
DB 345 TTCAAGCAAACTGGAAATAATCAATCAATTAATGCTTTCTTCCCAAGAGTTTTTAA 404

QY 1935 TGATATGCCAGCTTCCTAAATTTGGAGACAAAAG 1967
DB 405 TGATATGCCAGCTTCCTAAATTTGGAGACAAAAG 437

RESULT 14
ABV40245
ID ABV40245 standard; cDNA; 437 BP.
XX
AC ABV40245;
XX
DT 16-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 40236.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW Pharmacogenomic marker; Gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US005171.
XX
PR 17-FEB-2000; 2000US-0183319P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JE;
XX
DR WPI; 2001-662795/76.
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS Claim 1; Page 8129; 11750pp; English.

CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker

SQ Sequence 437 BP; 150 A; 63 C; 79 G; 145 T; 0 U; 0 Other;

Query Match 9.1%; Score 342; DB 5; Length 437;
Best Local Similarity 99.7%; Pred. No. 8.6e-109;
Matches 392; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1575 GTACTGGATATTTTGCATATAAATTCGAGTATAGTTCAAAAATTAATAGTTTTTAC 1634
DB 45 GTACTGGATATTTTGCATATAAATTCGAGTATAGTTCAAAAATTAATAGTTTTTAC 104

QY 1635 ATTGCTTTTCTGAGAGAGAAATTTGAAAGTGTCAAAAATAAAAAAGATGAAATGAAG 1694
DB 105 ATTGCTTTTCTGAGAGAGAAATTTGAAAGTGTCAAAAATAAAAAAGATGAAATGAAG 164

QY 1695 CATATATAATTTCTCAATTTTTCATTTTCTAGTCAACAGAGAAATCGAAGGATCTGTTC 1754
DB 165 CATATATAATTTCTCAATTTTTCATTTTCTAGTCAACAGAGAAATCGAAGGATCTGTTC 224

QY 1755 AAATATTAGTAAAAATTTGAAAAATAAATTTGCTTTATATTTTGTTCACACACTAGTT 1814
DB 225 AAATATTAGTAAAAATTTGAAAAATAAATTTGCTTTATATTTTGTTCACACACTAGTT 284

QY 1815 AATTTAACTGTGACTAGTTATCTCTACGAGGTGGATGTGTAGTTCTCGTTTAAAA 1874
DB 285 AATTTAACTGTGACTAGTTATCTCTACGAGGTGGATGTGTAGTTCTCGTTTAAAA 344

QY 1875 TTCAAGCAAACTGGAAATAATCAATCAATTAATGCTTTCTTCCCAAGAGTTTTTAA 1934
DB 345 TTCAAGCAAACTGGAAATAATCAATCAATTAATGCTTTCTTCCCAAGAGTTTTTAA 404

QY 1935 TGATATGCCAGCTTCCTAAATTTGGAGACAAAAG 1967
DB 405 TGATATGCCAGCTTCCTAAATTTGGAGACAAAAG 437

RESULT 15
ABV40577
ID ABV40577 standard; cDNA; 437 BP.
XX
AC ABV40577;
XX
DT 16-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 40568.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW Pharmacogenomic marker; Gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US005171.
XX
PR 17-FEB-2000; 2000US-0183319P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JE;
XX
DR WPI; 2001-662795/76.
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.

XX Claim 1; Page 8181; 11750pp; English.

CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;

CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
 CC determining whether prostate cancer has metastasized in a patient; (h)
 CC assessing the aggressiveness or indolence of prostate cancer in a patient
 CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
 XX
 SQ Sequence 437 BP; 150 A; 63 C; 79 G; 145 T; 0 U; 0 Other;

Query Match 9.1%; Score 342; DB 5; Length 437;
 Best Local Similarity 99.7%; Pred. No. 8.6e-109;
 Matches 392; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1575 GTACTGATATTTTGCATATAAATTCAGTATGCTTCAAAAATTAATAGTTTGGAC 1634
 Db |||||
 QY 45 GTACTGATATTTTGCATATAAATTCAGTATGCTTCAAAAATTAATAGTTTGGAC 104
 Db |||||
 QY 1635 ATTGGCTTTTCTGAGAGAGAAATGGAAGTGTCAAAAATAAAAAGATGAATGAAG 1694
 Db |||||
 QY 105 ATTGGCTTTTCTGAGAGAGAAATGGAAGTGTCAAAAATAAAAAGATGAATGAAG 164
 Db |||||
 QY 1695 CATATATAATTTGCAATTTTTCATTTTCTAGTCAACAGAGATCGAAGGATTCGTTC 1754
 Db |||||
 QY 165 CATATATAATTTGCAATTTTTCATTTTCTAGTCAACAGAGATCGAAGGATTCGTTC 224
 Db |||||
 QY 1755 AAATATTAGTAAAAATGAAATTAATCTGTGCTTATATTTTGTTCGAACACACTAGTT 1814
 Db |||||
 QY 225 AAATATTAGTAAAAATGAAATTAATCTGTGCTTATATTTTGTTCGAACACACTAGTT 284
 Db |||||
 QY 1815 AATTAACTGTGACTAGTTATCTCTACCGAAGTGGATGATGTTCTGGTTTAAAA 1874
 Db |||||
 QY 285 AATTAACTGTGACTAGTTATCTCTACCGAAGTGGATGATGTTCTGGTTTAAAA 344
 Db |||||
 QY 1875 TTCAAGCAACTGGAAATAATCCATCTAATATGCTTCTTCCCAAGAAATTTTAA 1934
 Db |||||
 QY 345 TTCAAGCAACTGGAAATAATCCATCTAATATGCTTCTTCCCAAGAAATTTTAA 404
 Db |||||
 QY 1935 TGATATGCCAGTCTCTAATTTGGAGACAAAAG 1967
 Db |||||
 QY 405 TGATATGCCAGTCTCTAATTTGGAGACAAAAG 437
 Db |||||

RESULT 16
 ABV00935
 ID ABV00935 standard; cDNA; 446 BP.
 XX
 AC ABV00935;
 XX
 DT 13-SEP-2002 (first entry)
 XX
 DE Human prostate expression marker cDNA 926.
 XX
 KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200160860-A2.
 XX
 PD 23-AUG-2001.
 XX
 PF 20-FEB-2001; 2001WO-US005171.
 XX
 PR 17-FEB-2000; 2000US-0183319P.
 PR 16-MAR-2000; 2000US-0189862P.
 PR 25-MAY-2000; 2000US-0207454P.
 PR 09-JUN-2000; 2000US-0211314P.
 PR 18-JUL-2000; 2000US-0219007P.
 PR 13-DEC-2000; 2000US-0255281P.
 XX
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX
 PI Schlegel R, Endege WO, Monahan JE;
 XX
 DR WPI; 2001-662795/76.

XX Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer.
 XX
 PS Claim 1; Page 246; 11750pp; English.
 XX
 CC The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for: (a) assessing whether
 CC a patient is afflicted with prostate cancer; (b) monitoring the
 CC progression of prostate cancer in a patient; (c) assessing the efficacy
 CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
 CC determining whether prostate cancer has metastasized in a patient; (h)
 CC assessing the aggressiveness or indolence of prostate cancer in a patient
 CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
 XX
 SQ Sequence 446 BP; 146 A; 68 C; 77 G; 152 T; 0 U; 3 Other;

Query Match 6.5%; Score 244; DB 5; Length 446;
 Best Local Similarity 99.4%; Pred. No. 7e-75;
 Matches 344; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1611 GTTCAAAAATTAATAGTTTTCACATTCGCTTTCTGAGAGAGAAATGAAAGTGTCA 1670
 Db |||||
 QY 80 GTTCAAAAATTAATAGTTTTCACATTCGCTTTCTGAGAGAGAAATGAAAGTGTCA 139
 Db |||||
 QY 1671 AAAATAAAAAAGATGAATGAAGCATATATAATTTGTCAATTTTCAATTTTCTAGTCA 1730
 Db |||||
 QY 140 AAAATAAAAAAGATGAATGAAGCATATATAATTTGTCAATTTTCAATTTTCTAGTCA 199
 Db |||||
 QY 1731 ACAGAGAAATCGAAGATTCGTTCAAATATTAGTAAAAATTCAAAATAAATCTGCTTA 1790
 Db |||||
 QY 200 ACAGAGAAATCGAAGATTCGTTCAAATATTAGTAAAAATTCAAAATAAATCTGCTTA 259
 Db |||||
 QY 1791 TATTTTGTTCGAACACACTAGTTAATTTAACTGTGACTAGTTATCTCTACCGAAGTG 1850
 Db |||||
 QY 260 TATTTTGTTCGAACACACTAGTTAATTTAACTGTGACTAGTTATCTCTACCGAAGTG 319
 Db |||||
 QY 1851 GATGTGATTTTCTGGTTTTTAAATTCAGCAAACTGGAATAATCCATCTAATTTATGC 1910
 Db |||||
 QY 320 GATGTGATTTTCTGGTTTTTAAATTCAGCAAACTGGAATAATCCATCTAATTTATGC 379
 Db |||||
 QY 1911 TTCTTTCCCAAGAGTTTTTTAATGATATGCCAGTCTCTCTAATTT 1956
 Db |||||
 QY 380 TTCTTTCCCAAGAGTTTTTTAATGATATGCCAGTCTCTCTAATTT 425
 Db |||||

RESULT 17
 ABV10104
 ID ABV10104 standard; cDNA; 404 BP.
 XX
 AC ABV10104;
 XX
 DT 13-SEP-2002 (first entry)
 XX
 DE Human prostate expression marker cDNA 10095.
 XX
 KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200160860-A2.
 XX
 PD 23-AUG-2001.
 XX
 PF 20-FEB-2001; 2001WO-US005171.
 XX
 DR 17-FEB-2000; 2000US-0183319P.

```
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX Schlegel R, Endege WO, Monahan JE;
XX WPI; 2001-662795/76.
XX Novel isolated nucleic acid molecule associated with cancerous state of
XX prostate cells and correlating with presence of prostate cancer, useful
XX for detecting presence of prostate cancer, stage of prostate cancer.
XX Claim 1; Page 1617-1618; 11750pp; English.
XX The invention relates to an isolated nucleic acid molecule (I) comprising
XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX specification or its complement. (I) is useful for: (a) assessing whether
XX a patient is afflicted with prostate cancer; (b) monitoring the
XX progression of prostate cancer in a patient; (c) assessing the efficacy
XX of a test compound to inhibit prostate cancer in a patient; (d) assessing
XX the efficacy of a therapy for inhibiting prostate cancer in a patient;
XX (e) selecting a composition for inhibiting prostate cancer in a patient;
XX (f) assessing the prostate cell carcinogenic potential of a compound; (g)
XX determining whether prostate cancer has metastasized in a patient; (h)
XX assessing the aggressiveness or indolence of prostate cancer in a patient
XX ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
XX Sequence 404 BP; 142 A; 54 C; 63 G; 144 T; 0 U; 1 Other;
XX
XX Query Match 5.7%; Score 213; DB 5; Length 404;
XX Best Local Similarity 99.2%; Pred. No. 3.8e-64;
XX Matches 363; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
XX
QY 1611 GTTCAAAAATTAAATAGTTTTTGACATTGGCTTTTCTGAGAGAGAAATGAAAGTGTAC 1670
DB 39 GTTCAAAAATTAAATAGTTTTTGACATTGGCTTTTCTGAGAGAGAAATGAAAGTGTAC 98
QY 1671 AAATTAATAAAGATGAATGAGCATATATATTTCAATTTTCTCAATTTTCTAGTCA 1730
DB 99 AAATTAATAAAGATGAATGAGCATATATATTTCAATTTTCTCAATTTTCTAGTCA 158
QY 1731 ACAGAGATCGAAGGATCTGTTCAAAATATTAGTAAATAATGAAATAAATCTGTGCTTA 1790
DB 159 ACAGAGATCGAAGGATCTGTTCAAAATATTAGTAAATAATGAAATAAATCTGTGCTTA 218
QY 1791 TATTTTGTTCGACACACACTAGTTAATTTAACTGTGACTAGTTATCTTACCGAAGTG 1850
DB 219 TATTTTGTTCGACACACACTAGTTAATTTAACTGTGACTAGTTATCTTACCGAAGTG 278
QY 1851 GATGTGTAGTTTCTGTTTAAATTTCAAGCAAACTGGAATAATCCATCTAATTATGC 1910
DB 279 GATGTGTAGTTTCTGTTTAAATTTCAAGCAAACTGGAATAATCCATCTAATTATGC 338
QY 1911 TTTCTTTCCCAAGAGATTTTAAATGATGATGCCAGCTTCTTAATTTGGAGACAAAGCCT 1970
DB 339 TTTCTTTCCCAAGAGATTTTAAATGATGATGCCAGCTTCTTAATTTGGAGACAAAGCCT 398
QY 1971 TAAATG 1976
DB 399 TAAATG 404
RESULT 18
ABV31276
ID ABV31276 standard; cDNA; 421 BP.
XX
AC ABV31276;
XX
DT 16-SEP-2002 (first entry)
```

```
XX Human prostate expression marker cDNA 31267.
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX pharmacogenomic marker; gene; ss.
XX Homo sapiens.
XX WO200160860-A2.
XX 23-AUG-2001.
XX 20-FEB-2001; 2001WO-US005171.
XX 17-FEB-2000; 2000US-0183319P.
XX 16-MAR-2000; 2000US-0189862P.
XX 25-MAY-2000; 2000US-0207454P.
XX 09-JUN-2000; 2000US-0211314P.
XX 18-JUL-2000; 2000US-0219007P.
XX 13-DEC-2000; 2000US-0255281P.
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX Schlegel R, Endege WO, Monahan JE;
XX WPI; 2001-662795/76.
XX Novel isolated nucleic acid molecule associated with cancerous state of
XX prostate cells and correlating with presence of prostate cancer, useful
XX for detecting presence of prostate cancer, stage of prostate cancer.
XX Claim 1; Page 6740; 11750pp; English.
XX The invention relates to an isolated nucleic acid molecule (I) comprising
XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX specification or its complement. (I) is useful for: (a) assessing whether
XX a patient is afflicted with prostate cancer; (b) monitoring the
XX progression of prostate cancer in a patient; (c) assessing the efficacy
XX of a test compound to inhibit prostate cancer in a patient; (d) assessing
XX the efficacy of a therapy for inhibiting prostate cancer in a patient;
XX (e) selecting a composition for inhibiting prostate cancer in a patient;
XX (f) assessing the prostate cell carcinogenic potential of a compound; (g)
XX determining whether prostate cancer has metastasized in a patient; (h)
XX assessing the aggressiveness or indolence of prostate cancer in a patient
XX ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
XX Sequence 421 BP; 143 A; 61 C; 71 G; 145 T; 0 U; 1 Other;
XX
XX Query Match 5.7%; Score 213; DB 5; Length 421;
XX Best Local Similarity 99.2%; Pred. No. 3.8e-64;
XX Matches 363; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
XX
QY 1611 GTTCAAAAATTAAATAGTTTTTGACATTGGCTTTTCTGAGAGAGAAATGAAAGTGTAC 1670
DB 56 GTTCAAAAATTAAATAGTTTTTGACATTGGCTTTTCTGAGAGAGAAATGAAAGTGTAC 115
QY 1671 AAATTAATAAAGATGAATGAGCATATATATTTCAATTTTCTCAATTTTCTAGTCA 1730
DB 116 AAATTAATAAAGATGAATGAGCATATATATTTCAATTTTCTCAATTTTCTAGTCA 175
QY 1731 ACAGAGATCGAAGGATCTGTTCAAAATATTAGTAAATAATGAAATAAATCTGTGCTTA 1790
DB 176 ACAGAGATCGAAGGATCTGTTCAAAATATTAGTAAATAATGAAATAAATCTGTGCTTA 235
QY 1791 TATTTTGTTCGACACACACTAGTTAATTTAACTGTGACTAGTTATCTTACCGAAGTG 1850
DB 236 TATTTTGTTCGACACACACTAGTTAATTTAACTGTGACTAGTTATCTTACCGAAGTG 295
QY 1851 GATGTGTAGTTTCTGTTTAAATTTCAAGCAAACTGGAATAATCCATCTAATTATGC 1910
DB 296 GATGTGTAGTTTCTGTTTAAATTTCAAGCAAACTGGAATAATCCATCTAATTATGC 355
QY 1911 TTTCTTTCCCAAGAGATTTTAAATGATGATGCCAGCTTCTTAATTTGGAGACAAAGCCT 1970
```

```
Db 356 TTTCTTCCCAAGAGTTTATGATATGCAGCTTCTTAATTTGGACAAAAGCT 415
QY 1971 TAATTG 1976
Db 416 TAATTG 421

RESULT 19
ABV09593
ID ABV09593 standard; cDNA; 359 BP.
XX AC ABV09593;
XX 13-SEP-2002 (first entry)
XX Human prostate expression marker cDNA 9584.
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX Homo sapiens.
XX WO200160860-A2.
XX 23-AUG-2001.
XX 20-FEB-2001; 2001WO-US005171.
XX 17-FEB-2000; 2000US-0183319P.
XX 16-MAR-2000; 2000US-0189862P.
XX 25-MAY-2000; 2000US-0207454P.
XX 09-JUN-2000; 2000US-0211314P.
XX 18-JUL-2000; 2000US-0219007P.
XX 13-DEC-2000; 2000US-0255281P.
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX Schlegel R, Endege WO, Monahan JE;
XX WPI; 2001-662795/76.
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX Claim 1; Page 1523; 11750pp; English.
XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
XX Sequence 359 BP; 128 A; 45 C; 57 G; 128 T; 0 U; 1 Other;
Query Match 5.4%; Score 204; DB 5; Length 359;
Best Local Similarity 100.0%; Pred. No. 5.1e-61;
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1729 CAACAGAGATCGAAGGATTCGTTCAAATATTAGTAAATTAAGAAATTAACCTTGCT 1788
Db 155 CAACAGAGATCGAAGGATTCGTTCAAATATTAGTAAATTAAGAAATTAACCTTGCT 214
QY 1789 TATATTTTGGCAACACACTAGTTAATTTAACTGTGACTAGTTATCTCTACCGAAG 1848
|||||
```

```
Db 215 TATATTTTGGTTCACACACTAGTTAATTTAACTGTGACTAGTTATCTCTACCGAAG 274
QY 1849 TCGATGTCTAGTTCTTGGTTTTTAAATTCAGCAAACTGGAAATTAATCCATCTAATTAT 1908
|||||
Db 275 TGGATGTCTAGTTCTTGGTTTTTAAATTCAGCAAACTGGAAATTAATCCATCTAATTAT 334
QY 1909 GCTTCTTCTTCCCAAGAGTTTATTTT 1932
Db 335 GCTTCTTCTTCCCAAGAGTTTATTTT 358

RESULT 20
ABV00424
ID ABV00424 standard; cDNA; 400 BP.
XX AC ABV00424;
XX 13-SEP-2002 (first entry)
XX Human prostate expression marker cDNA 415.
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX Homo sapiens.
XX WO200160860-A2.
XX 23-AUG-2001.
XX 20-FEB-2001; 2001WO-US005171.
XX 17-FEB-2000; 2000US-0183319P.
XX 16-MAR-2000; 2000US-0189862P.
XX 25-MAY-2000; 2000US-0207454P.
XX 09-JUN-2000; 2000US-0211314P.
XX 18-JUL-2000; 2000US-0219007P.
XX 13-DEC-2000; 2000US-0255281P.
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX Schlegel R, Endege WO, Monahan JE;
XX WPI; 2001-662795/76.
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX Claim 1; Page 171; 11750pp; English.
XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
XX Sequence 400 BP; 134 A; 59 C; 73 G; 134 T; 0 U; 0 Other;
Query Match 5.4%; Score 203; DB 5; Length 400;
Best Local Similarity 100.0%; Pred. No. 1.1e-60;
Matches 203; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1729 CAACAGAGATCGAAGGATTCGTTCAAATATTAGTAAATTAAGAAATTAACCTTGCT 1788
Db 198 CAACAGAGATCGAAGGATTCGTTCAAATATTAGTAAATTAAGAAATTAACCTTGCT 257
|||||
```

QY 1789 TATATTTTGTGCAACACACTAGTAAATTAACCTGTGACTAGTATCTCTACCGAAGG 1848
DB 258 TATATTTTGTGCAACACACTAGTAAATTAACCTGTGACTAGTATCTCTACCGAAGG 317
QY 1849 TGGATGTGAGTCTTCTGTTTAAATTCAGCAAACTGGAAATATCCATCTAATTAT 1908
DB 318 TGGATGTGAGTCTTCTGTTTAAATTCAGCAAACTGGAAATATCCATCTAATTAT 377
QY 1909 GCTTTCTTTCCCAAGAGTTTTT 1931
DB 378 GCTTTCTTTCCCAAGAGTTTTT 400

RESULT 21
ABV39737
ID ABV39737 standard; cDNA; 401 BP.
XX
AC ABV39737;
XX
DT 16-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 39728.
XX
DE Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
KW Homo sapiens.
XX
OS WO200160860-A2.
XX
PN 23-AUG-2001.
XX
PD 20-FEB-2001; 2001WO-US005171.
XX
PF 17-FEB-2000; 2000US-0183319P.
XX
PR 16-MAR-2000; 2000US-0189862P.
XX
PR 25-MAY-2000; 2000US-0207454P.
XX
PR 09-JUN-2000; 2000US-0211314P.
XX
PR 18-JUL-2000; 2000US-0219007P.
XX
PR 13-DEC-2000; 2000US-0255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JE;
XX
DR WPI; 2001-662795/76.
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS Claim 1; Page 8048; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the efficacy
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 401 BP; 136 A; 58 C; 73 G; 134 T; 0 U; 0 Other;

Query Match 5.4%; Score 203; DB 5; Length 401;
Best Local Similarity 100.0%; Pred. No. 1.1e-60; Indels 0; Gaps 0;
Matches 203; Conservative 0; Mismatches 0;

QY 1729 CAACAGAGATCGAAGGATTCGTTCAAATATAGTAAATTAAGAAATAAATCTGTGCT 1788
DB 199 CAACAGAGATCGAAGGATTCGTTCAAATATAGTAAATTAAGAAATAAATCTGTGCT 258
QY 1789 TATATTTTGTGCAACACACTAGTAAATTAACCTGTGACTAGTATCTCTACCGAAGG 1848
DB 259 TATATTTTGTGCAACACACTAGTAAATTAACCTGTGACTAGTATCTCTACCGAAGG 318
QY 1849 TGGATGTGAGTCTTCTGTTTAAATTCAGCAAACTGGAAATATCCATCTAATTAT 1908
DB 319 TGGATGTGAGTCTTCTGTTTAAATTCAGCAAACTGGAAATATCCATCTAATTAT 378
QY 1909 GCTTTCTTTCCCAAGAGTTTTT 1931
DB 379 GCTTTCTTTCCCAAGAGTTTTT 401

RESULT 22
ABV30769
ID ABV30769 standard; cDNA; 401 BP.
XX
AC ABV30769;
XX
DT 16-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 30760.
XX
DE Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
KW Homo sapiens.
XX
OS WO200160860-A2.
XX
PN 23-AUG-2001.
XX
PD 20-FEB-2001; 2001WO-US005171.
XX
PF 17-FEB-2000; 2000US-0183319P.
XX
PR 16-MAR-2000; 2000US-0189862P.
XX
PR 25-MAY-2000; 2000US-0207454P.
XX
PR 09-JUN-2000; 2000US-0211314P.
XX
PR 18-JUL-2000; 2000US-0219007P.
XX
PR 13-DEC-2000; 2000US-0255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JE;
XX
DR WPI; 2001-662795/76.
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS Claim 1; Page 6663-6664; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the efficacy
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 401 BP; 136 A; 58 C; 73 G; 134 T; 0 U; 0 Other;

Query Match 5.4%; Score 203; DB 5; Length 401;

```
Best Local Similarity 100.0%; Pred. No. 1.1e-60;
Matches 203; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1729 CAACAGAGATCGAAGGATTCCTGTTCAAAATATTAGTAAATAATTTGAAAATAAACTTGTGCT 1788
DB 199 CAACAGAGATCGAAGGATTCCTGTTCAAAATATTAGTAAATAATTTGAAAATAAACTTGTGCT 258
QY 1789 TATATTTTGTTCGAACACACTAGTTAAATTTAACTGTGACTAGTTATCTCTACCGAAGG 1848
DB 259 TATATTTTGTTCGAACACACTAGTTAAATTTAACTGTGACTAGTTATCTCTACCGAAGG 318
QY 1849 TGGATGTGTAGTTCTGTTTAAATTTCAAGCAAACTGGAATAATTCATCTAATTAT 1908
DB 319 TGGATGTGTAGTTCTGTTTAAATTTCAAGCAAACTGGAATAATTCATCTAATTAT 378
QY 1909 GCTTTCTTTCCCAAGAGTTTTT 1931
DB 379 GCTTTCTTTCCCAAGAGTTTTT 401

RESULT 23
ABV39958
ID ABV39958 standard; cDNA; 401 BP.
XX AC ABV39958;
XX DT 16-SEP-2002 (first entry)
XX DE Human prostate expression marker cDNA 39949.
XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX KW pharmacogenomic marker; gene; ss.
XX OS Homo sapiens.
XX PN WO200160860-A2.
XX PD 23-AUG-2001.
XX PF 20-FEB-2001; 2001WO-US005171.
XX PR 17-FEB-2000; 2000US-0183319P.
XX PR 16-MAR-2000; 2000US-0189862P.
XX PR 25-MAY-2000; 2000US-0207454P.
XX PR 09-JUN-2000; 2000US-0211314P.
XX PR 18-JUL-2000; 2000US-0219007P.
XX PR 13-DEC-2000; 2000US-0255281P.
XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX PI Schlegel R, Endege WO, Monahan JE;
XX DR WPI; 2001-662795/76.
XX PT Novel isolated nucleic acid molecule associated with cancerous state of
XX PT prostate cells and correlating with presence of prostate cancer, useful
XX PT for detecting presence of prostate cancer, stage of prostate cancer.
XX PS Claim 1; Page 8083; 11750pp; English.
XX CC The invention relates to an isolated nucleic acid molecule (I) comprising
XX CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX CC specification or its complement. (I) is useful for: (a) assessing whether
XX CC a patient is afflicted with prostate cancer; (b) monitoring the
XX CC progression of prostate cancer in a patient; (c) assessing the efficacy
XX CC of a test compound to inhibit prostate cancer in a patient;
XX CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
XX CC (e) assessing the prostate cell carcinogenic potential of a compound; (g)
XX CC determining whether prostate cancer has metastasized in a patient; (h)
XX CC assessing the aggressiveness or indolence of prostate cancer in a patient
XX CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker

SQ Sequence 401 BP; 136 A; 58 C; 73 G; 134 T; 0 U; 0 Other;
Query Match 5.4%; Score 203; DB 5; Length 401;
Best Local Similarity 100.0%; Pred. No. 1.1e-60;
Matches 203; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1729 CAACAGAGATCGAAGGATTCCTGTTCAAAATATTAGTAAATAATTTGAAAATAAACTTGTGCT 1788
DB 199 CAACAGAGATCGAAGGATTCCTGTTCAAAATATTAGTAAATAATTTGAAAATAAACTTGTGCT 258
QY 1789 TATATTTTGTTCGAACACACTAGTTAAATTTAACTGTGACTAGTTATCTCTACCGAAGG 1848
DB 259 TATATTTTGTTCGAACACACTAGTTAAATTTAACTGTGACTAGTTATCTCTACCGAAGG 318
QY 1849 TGGATGTGTAGTTCTGTTTAAATTTCAAGCAAACTGGAATAATTCATCTAATTAT 1908
DB 319 TGGATGTGTAGTTCTGTTTAAATTTCAAGCAAACTGGAATAATTCATCTAATTAT 378
QY 1909 GCTTTCTTTCCCAAGAGTTTTT 1931
DB 379 GCTTTCTTTCCCAAGAGTTTTT 401

RESULT 24
ABV30990
ID ABV30990 standard; cDNA; 296 BP.
XX AC ABV30990;
XX DT 16-SEP-2002 (first entry)
XX DE Human prostate expression marker cDNA 30981.
XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX KW pharmacogenomic marker; gene; ss.
XX OS Homo sapiens.
XX PN WO200160860-A2.
XX PD 23-AUG-2001.
XX PF 20-FEB-2001; 2001WO-US005171.
XX PR 17-FEB-2000; 2000US-0183319P.
XX PR 16-MAR-2000; 2000US-0189862P.
XX PR 25-MAY-2000; 2000US-0207454P.
XX PR 09-JUN-2000; 2000US-0211314P.
XX PR 18-JUL-2000; 2000US-0219007P.
XX PR 13-DEC-2000; 2000US-0255281P.
XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX PI Schlegel R, Endege WO, Monahan JE;
XX DR WPI; 2001-662795/76.
XX PT Novel isolated nucleic acid molecule associated with cancerous state of
XX PT prostate cells and correlating with presence of prostate cancer, useful
XX PT for detecting presence of prostate cancer, stage of prostate cancer.
XX PS Claim 1; Page 6697; 11750pp; English.
XX CC The invention relates to an isolated nucleic acid molecule (I) comprising
XX CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX CC specification or its complement. (I) is useful for: (a) assessing whether
XX CC a patient is afflicted with prostate cancer; (b) monitoring the
XX CC progression of prostate cancer in a patient; (c) assessing the efficacy
XX CC of a test compound to inhibit prostate cancer in a patient;
XX CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
XX CC (e) assessing the prostate cell carcinogenic potential of a compound; (g)
XX CC determining whether prostate cancer has metastasized in a patient; (h)
XX CC assessing the aggressiveness or indolence of prostate cancer in a patient
XX CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
```



```
Db 193 TAAAAATTGAAATAAATCTGTCTATATTGTTGCAACACACTAGTAAATTAACC 252
QY 1824 TG 1825
Db 253 TG 254

RESULT 27
ABV56428
ID ABV56428 standard; cDNA; 381 BP.
XX
AC ABV56428;
XX
DT 17-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 56419.
XX
DE Human prostate expression marker cDNA 56419.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US005171.
XX
PR 17-FEB-2000; 2000US-0183319P.
XX
PR 16-MAR-2000; 2000US-0189862P.
XX
PR 25-MAY-2000; 2000US-0207454P.
XX
PR 09-JUN-2000; 2000US-0211314P.
XX
PR 18-JUL-2000; 2000US-0219007P.
XX
PR 13-DEC-2000; 2000US-0255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JE;
XX
WPI; 2001-662795/76.
XX
Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS Claim 1; Page 10887; 11750pp; English.
XX
The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 381 BP; 165 A; 67 C; 70 G; 78 T; 0 U; 1 Other;

Query Match 1.4%; Score 52; DB 5; Length 381;
Best Local Similarity 100.0%; Pred. No. 2e-08;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3711 TATAAATGAAAAAATAAATCTGTCTATATTGTTGCAACACACTAGTAAATTAACC 3762
Db 230 TATAAATGAAAAAATAAATCTGTCTATATTGTTGCAACACACTAGTAAATTAACC 281

RESULT 28
```

```
AAI90783
ID AAI90783 standard; cDNA; 345 BP.
XX
AC AAI90783;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 10843.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation; ss.
XX
OS Homo sapiens.
XX
PN WO200164835-A2.
XX
PD 07-SEP-2001.
XX
PF 26-FEB-2001; 2001WO-US004927.
XX
PR 28-FEB-2000; 2000US-00515126.
XX
PR 18-MAY-2000; 2000US-00577409.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
WPI; 2001-514838/56.
XX
P-PSDB; AAO10852.
XX
Isolated nucleic acids and polypeptides, useful for preventing diagnosing
PT and treating e.g. leukemia, inflammation and immune disorders.
XX
PS Claim 1; SEQ ID NO 10843; 1399pp + Sequence Listing; English.
XX
The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 345 BP; 142 A; 54 C; 69 G; 79 T; 0 U; 1 Other;

Query Match 1.3%; Score 50; DB 4; Length 345;
Best Local Similarity 100.0%; Pred. No. 1e-07;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAAAAATAAATCTGTCTATATTGTTGCAACACACTAGTAAATTAACC 3762
Db 278 TAAATGAAAAAATAAATCTGTCTATATTGTTGCAACACACTAGTAAATTAACC 327

RESULT 29
ABV57393
ID ABV57393 standard; cDNA; 424 BP.
XX
AC ABV57393;
XX
DT 17-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 57384.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
```

KW pharmacogenomic marker; gene; ss.

OS Homo sapiens.

XX WO200160860-A2.

FN 23-AUG-2001.

PD 20-FEB-2001; 2001WO-US005171.

XX 17-FEB-2000; 2000US-0183319P.

XX 16-MAR-2000; 2000US-0189862P.

PR 25-MAY-2000; 2000US-0207454P.

PR 09-JUN-2000; 2000US-0211314P.

PR 18-JUL-2000; 2000US-0219007P.

PR 13-DEC-2000; 2000US-0255281P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Schlegel R, Endege WO, Monahan JB;

PI WPI; 2001-662795/76.

XX Novel isolated nucleic acid molecule associated with cancerous state of

XX prostate cells and correlating with presence of prostate cancer, useful

XX for detecting presence of prostate cancer, stage of prostate cancer.

XX Claim 1; Page 11041; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker

XX Sequence 424 BP; 100 A; 114 C; 73 G; 136 T; 0 U; 1 Other;

Query Match 1.3%; Score 50; DB 5; Length 424;

Best Local Similarity 100.0%; Pred. No. 9.7e-08;

Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAA 3762

Db 47 TAAATGAAA 96

RESULT 30

AAI83050

ID AAI83050 standard; cDNA; 449 BP.

XX AA183050;

XX 06-NOV-2001 (first entry)

XX Human polynucleotide SEQ ID NO 3110.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;

XX vaccine; peptide therapy; stem cell growth factor; haematopoiesis;

XX tissue growth factor; immunomodulatory; cancer; leukaemia;

XX nervous system disorders; arthritis; inflammation; ss.

OS Homo sapiens.

XX WO200164835-A2.

XX 07-SEP-2001.

XX

PF

XX

PR

PR

XX

XX

PA

XX

XX

PI

XX

DR

DR

XX

XX

PT

PT

XX

XX

PS

XX

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

26-FEB-2001; 2001WO-US004927.

28-FEB-2000; 2000US-00515126.

18-MAY-2000; 2000US-00577409.

(HYSE-) HYSEQ INC.

Tang YT, Liu C, Drmanac RT;

WPI; 2001-514838/56.

P-PSDB; AA003119.

Isolated nucleic acids and polypeptides, useful for preventing diagnosing

and treating e.g. leukemia, inflammation and immune disorders.

Claim 1; SEQ ID NO 3110; 1399pp + Sequence Listing; English.

The invention relates to human polynucleotides (AAI79941-AAI93841) and

the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to

cytokine, cell proliferation or cell differentiation or which may induce

production of other cytokines in other cell populations. The

polynucleotides and polypeptides are useful in gene therapy, vaccines or

polynucleotides and polypeptides have various cytokine-like activities,

e.g. stem cell growth factor activity, haematopoiesis regulating

activity, tissue growth factor activity, immunomodulatory activity and

activin/inhibin activity and may be useful in the diagnosis and/or

treatment of cancer, leukaemia, nervous system disorders, arthritis and

inflammation. Note: The sequence data for this patent did not form part

of the printed specification, but was obtained in electronic format

directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 449 BP; 184 A; 55 C; 100 G; 109 T; 0 U; 1 Other;

Query Match 1.3%; Score 50; DB 4; Length 449;

Best Local Similarity 100.0%; Pred. No. 9.5e-08;

Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAA 3762

Db 125 TAAATGAAA 174

RESULT 31

AAF15960

ID AAF15960 standard; cDNA; 1702 BP.

XX AAF15960;

XX 13-MAR-2001 (first entry)

XX Human prostate cancer antigen nucleotide sequence SEQ ID NO:395.

XX Human; prostate cancer; prostate cancer antigen; detection; diagnosis;

XX neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular;

XX vulnary; gastrointestinal; nephrotropic; antiinfective; gynaecological;

XX antibacterial; gene therapy; neural; immune; reproductive; renal;

XX gastrointestinal; pulmonary; cardiovascular; proliferative disorder;

XX wound; infectious disease; ss.

XX Homo sapiens.

XX WO200055174-A1.

XX 21-SEP-2000.

XX 08-MAR-2000; 2000WO-US005988.

XX 12-MAR-1999; 99US-0124270P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX (ROSE/) ROSEN C A.

XX

XX

XX

XX

XX

XX

XX

XX

XX

```

PI Rosen CA, Ruben SM;
XX
DR WPI: 2000-587513/55.
DR P-PSDB; AAB56757.
XX
PT Prostate cancer associated gene sequences, referred to as prostate cancer
PT antigens, useful for treatment, prevention, and diagnosis of disorders
PT such as prostate cancer.
XX
PS Claim 1; Page 885; 2338pp; English.
XX
CC AAF1566 to AAF16505 encode the human prostate cancer associated
CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
CC The prostate cancer antigens can have neuroprotective, cytostatic,
CC cardioactive, immunomodulatory, muscular, vulnery, gastrointestinal,
CC nephrotropic, antiinfective, gynaecological and antibacterial activities,
CC and can be used in gene therapy. The prostate cancer antigen
CC polynucleotides may be used for detection of prostate cancer, chromosome
CC identification, as chromosome markers, and for numerous other diagnostic
CC or research purposes. The prostate cancer antigens may be used to treat
CC disorders such as neural, immune, muscular, reproductive,
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
CC AAB57303 represent sequences used in the exemplification of the present
CC invention
XX
SQ Sequence 1702 BP; 608 A; 284 C; 302 G; 500 T; 0 U; 8 Other;
Query Match 1.3%; Score 50; DB 3; Length 1702;
Best Local Similarity 100.0%; Pred. No. 7.3e-08; Indels 0; Gaps 0;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3713 TAAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 1588 TAAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1637

RESULT 32
ADD18824
ID ADD18824 standard; DNA; 1820 BP.
XX
AC ADD18824;
XX
DT 15-JAN-2004 (first entry)
XX
DE Human disease related protein DNA sequence SeqID256.
XX
KW human; disease state; cytostatic; antiinflammatory; ophthalmological;
KW antiarteriosclerotic; vulnery; gene therapy;
KW hypoxia-regulated condition; tumorigenesis; angiogenesis; apoptosis;
KW inflammation; erythropoiesis; glycolysis; gluconeogenesis;
KW glucose transportation; catecholamine synthesis; iron transport;
KW nitric oxide synthesis; cancer; ischaemic condition; reperfusion injury;
KW retinopathy; neonatal stress; pre-eclampsia; atherosclerosis;
KW inflammatory condition; wound healing; gene; ds.
XX
OS Homo sapiens.
XX
PN WO2003018621-A2.
XX
PD 06-MAR-2003.
XX
PF 23-AUG-2002; 2002WO-GB003892.
XX
PR 23-AUG-2001; 2001GB-00020558.
XX
PR 05-OCT-2001; 2001GB-00024037.
XX
PA (OXFO-) OXFORD BIOMEDICA UK LTD.
XX
PI Kingsman SM, White J, Ward NR, Harris RA, Naylor S, Mundy CR;
DR WPI: 2003-290046/28.
DR P-PSDB; ADD18823.
XX

```

```

XX
PT New substantially purified polypeptide, useful for diagnosing or treating
PT a hypoxia-regulated condition, such as cancer, ischemia, reperfusion
PT injury, retinopathy, pre-eclampsia, atherosclerosis, inflammation, or
PT wound healing.
XX
PS Claim 27; SEQ ID NO 256; 424pp; English.
XX
CC This invention relates to novel human genes and gene product which are
CC implicated in certain disease states. Compounds which modulate the
CC proteins of the invention may have cytostatic, antiinflammatory,
CC ophthalmological, antiarteriosclerotic or vulnery activities. The
CC sequences of the invention may be useful for gene therapy. The invention
CC may be useful for diagnosing or treating a hypoxia-regulated condition,
CC such as tumorigenesis, angiogenesis, apoptosis, inflammation,
CC erythropoiesis, or the biological response to hypoxia conditions
CC including processes such as glycolysis, gluconeogenesis, glucose
CC transportation, catecholamine synthesis, iron transport or nitric oxide
CC synthesis. The disease includes cancer, ischaemic conditions, reperfusion
CC injury, retinopathy, neonatal stress, pre-eclampsia, atherosclerosis,
CC inflammatory conditions or wound healing. The present sequence is that of
CC a disease related protein encoding DNA sequence of the invention.
XX
SQ Sequence 1820 BP; 451 A; 485 C; 472 G; 412 T; 0 U; 0 Other;
Query Match 1.3%; Score 50; DB 9; Length 1820;
Best Local Similarity 100.0%; Pred. No. 7.3e-08; Indels 0; Gaps 0;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3713 TAAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 1719 TAAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1768

RESULT 33
ACF12861
ID ACF12861 standard; cDNA; 3602 BP.
XX
AC ACF12861;
XX
DT 10-SEP-2003 (first entry)
XX
DE Human cervical cancer cell marker encoding cDNA SEQ ID NO:67.
XX
KW Human; cervical cancer; cervical cancer marker; cancer therapy;
KW detection; gene therapy; vaccine; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2002101075-A2.
XX
PD 19-DEC-2002.
XX
PF 12-JUN-2002; 2002WO-US018638.
XX
PR 13-JUN-2001; 2001US-0298155P.
XX
PR 13-JUN-2001; 2001US-0298159P.
XX
PR 14-NOV-2001; 2001US-0335936P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Schlegel R, Chen Y, Zhao X, Monahan JE, Kamatkar S;
PI Gannavarapu M, Glatt K, Hoersch S;
XX
XX WPI: 2003-156967/15.
XX
PT New isolated nucleic acid molecule useful for detecting, characterizing,
PT preventing and treating human cervical cancers, in various prognostic and
PT diagnostic assays, in pharmacogenomics and in monitoring clinical trials.
XX
PS Claim 1; Page 218-219; 386pp; English.
XX
CC ACF12828 to ACF12947 encode the human cervical cancer marker proteins (I)

```

CC given in ABR92047 to ABR92164. A higher level of expression of (I) than
CC normal indicates the presence of cervical cancer. Also described: (1) a
CC vector (II) containing (I); (2) a host cell (III) containing (I); and (3)
CC assessing (MI) whether a patient is afflicted with cervical cancer,
CC comprising comparing the level of expression of a marker in a patient's
CC sample, and the normal level of expression of the marker in a control non
CC -cervical cancer sample, where a significant increase in the level of
CC expression of the marker in the patient's sample relative to that in the
CC control sample is an indication that the patient is afflicted with
CC cervical cancer. (I) has cytostatic activity, and can be used in gene
CC therapy and in vaccines. (I) is useful in detecting, characterising,
CC preventing and treating human cervical cancers. (I) may also be used in
CC various prognostic and diagnostic assays, pharmacogenomics and in
CC monitoring clinical trials

XX Sequence 3602 BP; 790 A; 1062 C; 1017 G; 730 T; 0 U; 3 Other;

Query Match 1.3%; Score 50; DB 7; Length 3602;
Best Local Similarity 100.0%; Pred. No. 6.3e-08;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAAAA 3762
DB 3501 TAAATGAAAAA 3550

RESULT 34
AA87412
ID AA87412 standard; cDNA; 3607 BP.

AC AA87412;

DT 08-OCT-1999 (first entry)

DE Hepatocellular carcinoma marker gene L4 cDNA.

XX Hepatocellular carcinoma; tumour; cancer; diagnosis; marker; probe;
KW human; hepatitis B virus; x antigen; HBxAg; L4 gene; ss.

XX Homo sapiens.

XX WO9939200-A1.

XX 05-AUG-1999.

XX 27-JAN-1999; 99WO-US001894.

XX 29-JAN-1998; 98US-0072938P.

XX (UYJE-) UNIV JEFFERSON THOMAS.

XX Feitelson MA;

XX WPI; 1999-469371/39.

XX New method for diagnosing hepatocellular carcinoma.

XX Claim 3; Page 43-44; 52pp; English.

XX This sequence represents a gene, designated L4, whose expression is
XX activated in human hepatoblastoma HepG2 cells in the presence of
XX hepatitis B virus x antigen (HBxAg) compared to expression in HepG2 cells
XX in the absence of HBxAg. The full-length cDNA was obtained from HepG2
XX cells using 5' and 3' RACE. No sequence homology was found between L4 and
XX any previously known sequence. Synthetic peptides that represent probable
XX antigenic determinants of the L4 protein are provided in AAY06538 and
XX AAY06539. L4 is 1 of 10 genes (see AA87402-11) that were initially
XX identified by PCR select cDNA subtraction as showing differential
XX expression in HBxAg(+) and HBxAg(-) cells. It can be used as a molecular
XX marker for hepatocellular carcinoma (HCC). In a claimed method for
XX detecting HCC, a liver tissue sample is obtained from a patient, and the
XX level of expression of 1 or more marker genes, such as L4, in the sample
XX is assessed. A reduction in the level of expression of the marker genes

CC in the sample as compared to the expression level in noncancerous liver
CC tissue is indicative of HCC. The method allows screening for the risk of
CC disease and early diagnosis before tumours develop

SQ Sequence 3607 BP; 875 A; 938 C; 1058 G; 736 T; 0 U; 0 Other;

Query Match 1.3%; Score 50; DB 2; Length 3607;

Best Local Similarity 100.0%; Pred. No. 6.3e-08;

Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAAAA 3762

DB 3531 TAAATGAAAAA 3580

RESULT 35

ABV48235

ID ABV48235 standard; cDNA; 265 BP.

AC ABV48235;

DT 17-SEP-2002 (first entry)

DE Human prostate expression marker CDNA 48226.

XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.

XX Homo sapiens.

XX WO200160860-A2.

XX 23-AUG-2001.

XX 20-FEB-2001; 2001WO-US005171.

XX 17-FEB-2000; 2000US-0183319P.

XX 16-MAR-2000; 2000US-0189862P.

XX 25-MAY-2000; 2000US-0207454P.

XX 09-JUN-2000; 2000US-0211314P.

XX 18-JUL-2000; 2000US-0219007P.

XX 13-DEC-2000; 2000US-0255281P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Schlegel R, Endege WO, Monahan JE;
XX WPI; 2001-652795/75.

XX Claim 1; Page 9463; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising
XX a nucleotide sequence given in Tables 1-9 (ABV0010-ABV62213) of the
XX specification or its complement. (I) is useful for: (a) assessing whether
XX a patient is afflicted with prostate cancer; (b) monitoring the
XX progression of prostate cancer in a patient; (c) assessing the efficacy
XX of a test compound to inhibit prostate cancer in a patient; (d) assessing
XX the efficacy of a therapy for inhibiting prostate cancer in a patient;
XX (e) selecting a composition for inhibiting prostate cancer in a patient;
XX (f) assessing the prostate cell carcinogenic potential of a compound; (g)
XX determining whether prostate cancer has metastasized in a patient; (h)
XX assessing the aggressiveness or indolence of prostate cancer in a patient
XX ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker

XX Sequence 265 BP; 159 A; 27 C; 31 G; 48 T; 0 U; 0 Other;

Query Match 1.3%; Score 49; DB 5; Length 265;

Best Local Similarity 100.0%; Pred. No. 2.3e-07;

Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


```
XX OS Homo sapiens.
XX PN WO200268645-A2.
XX PD 06-SEP-2002.
XX PF 20-NOV-2001; 2001WO-US045151.
XX PR 20-NOV-2000; 2000US-0249992P.
XX PA (DIAD-) DIADEXUS INC.
XX PI Salceda S, Macina RA, Recipon H, Cafferkey R, Sun Y, Liu C;
XX PI Turner LR;
XX WPI; 2002-713379/77.
XX DR New breast specific genes and proteins, useful in gene therapy or as
XX PT vaccines for treating breast cancer or non-cancerous breast diseases, as
XX PT well as for diagnosing, monitoring or staging these diseases.
XX PS Claim 1; Page 198-199; 277pp; English.
XX CC The invention relates to a novel isolated breast specific nucleic acid
XX CC molecule. The polypeptides of the invention have cytostatic activity. The
XX CC novel nucleic acids and polypeptides may have a use in gene therapy, and
XX CC as vaccines. The breast specific nucleic acid and polypeptide are useful
XX CC for diagnosing and monitoring the presence and metastases of lung cancer
XX CC in a patient. The antibody that specifically binds to the breast specific
XX CC polypeptide is useful for determining the presence of a breast specific
XX CC protein in a sample, as well as for treating a patient with breast
XX CC cancer, particularly by inducing an immune response against the breast
XX CC cancer cell expressing the breast specific nucleic acid molecule or
XX CC polypeptide. In particular, these breast specific genes and proteins are
XX CC useful for identifying, diagnosing, monitoring, staging, imaging and
XX CC treating breast cancer and non-cancerous disease states in breast tissue.
XX CC These are also useful in gene therapy, production of transgenic animals
XX CC and cells, and in the production of engineered breast tissue for
XX CC treatment and research. The sequences shown in AB277987-AB278101
XX CC represent the novel human breast specific nucleic acid molecules of the
XX CC invention
XX SQ Sequence 476 BP; 116 A; 88 C; 90 G; 182 T; 0 U; 0 Other;

Query Match 1.3%; Score 49; DB 6; Length 476;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
DB 126 AAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 78

RESULT 39
ABV58273
ID ABV58273 standard; cDNA; 497 BP.
XX AC ABV58273;
XX DT 13-SEP-2002 (first entry)
XX DE Human prostate expression marker cDNA 58264.
XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX KW pharmacogenomic marker; gene; ss.
XX OS Homo sapiens.
XX PN WO200160860-A2.
XX PD 23-AUG-2001.
```

```
PF 20-FEB-2001; 2001WO-US005171.
XX 17-FEB-2000; 2000US-0183319P.
XX PR 16-MAR-2000; 2000US-0189862P.
XX PR 25-MAY-2000; 2000US-0207454P.
XX PR 09-JUN-2000; 2000US-0211314P.
XX PR 18-JUL-2000; 2000US-0219007P.
XX PR 13-DEC-2000; 2000US-0255281P.
XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX PI Schlegel R, Endege WO, Monahan JB;
XX PI WPI; 2001-662795/76.
XX DR Novel isolated nucleic acid molecule associated with cancerous state of
XX DR prostate cells and correlating with presence of prostate cancer, useful
XX PT for detecting presence of prostate cancer, stage of prostate cancer.
XX PS Claim 1; Page 11189; 11750pp; English.
XX CC The invention relates to an isolated nucleic acid molecule (I) comprising
XX CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX CC specification or its complement. (I) is useful for: (a) assessing whether
XX CC a patient is afflicted with prostate cancer; (b) monitoring the
XX CC progression of prostate cancer in a patient; (c) assessing the efficacy
XX CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
XX CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
XX CC (e) selecting a composition for inhibiting prostate cancer in a patient;
XX CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
XX CC determining whether prostate cancer has metastasized in a patient; (h)
XX CC assessing the aggressiveness or indolence of prostate cancer in a patient
XX CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX SQ Sequence 497 BP; 235 A; 77 C; 69 G; 115 T; 0 U; 1 Other;

Query Match 1.3%; Score 49; DB 5; Length 497;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
DB 252 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 300

RESULT 40
AAT72173
ID AAT72173 standard; cDNA to mRNA; 882 BP.
XX AC AAT72173;
XX DT 25-FEB-1998 (first entry)
XX DE Alzheimer's disease DNA sequence from plasmid pGCS1180.
XX KW Human; brain; Alzheimer's disease; diagnosis; antibody; expression; ds.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT CDS 97..198
XX FT /*tag= a
XX FT /note= "protein encoded by DNA from the human brain which
XX FT has different expression from the normal protein when in
XX FT an Alzheimer's disease patient"
XX PN WO9721807-A1.
XX XX 19-JUN-1997.
XX PD 12-DEC-1996; 96WO-JP003630.
XX PF 12-DEC-1995; 95JP-00322745.
XX PR 12-DEC-1995; 95JP-00322745.
```

XX (KYOW) KYOWA HAKKO KOGYO KK.
 XX Ruga T, Nakagawa S, Sakaki Y, Zhao N, Hashida H;
 XX WPI; 1997-332779/30.
 DR P-PSDB; AAW19950.
 XX DNA sequences with varied expression in normal and Alzheimer patients -
 PT used for diagnosis and treatment of Alzheimer's disease.
 XX Claim 3; Page 48-49; 90pp; Japanese.
 XX The present sequence represents a DNA fragment which is complementary to
 CC mRNA sequences whose level of expression in the brains of Alzheimer's
 CC disease patients varies from normal individuals. The present sequence
 CC represents a specifically claimed DNA fragment in which the expression is
 CC augmented to at least twice that in normal cases. The sequence, and
 CC antibodies to the proteins encoded by the sequence, may be used in the
 CC diagnosis, treatment and investigation of Alzheimer's disease
 XX Sequence 882 BP; 306 A; 134 C; 142 G; 300 T; 0 U; 0 Other;
 SQ
 Query Match 1.3%; Score 49; DB 2; Length 882;
 Best Local Similarity 100.0%; Pred. NO. 1.9e-07;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 QY 3714 AATGAAAAA 3762
 DB 825 AATGAAAAA 873
 RESULT 41
 AAC79727
 ID AAC79727 standard; cDNA; 903 BP.
 AC AAC79727;
 XX
 XX 12-FEB-2001 (first entry)
 XX Human secreted protein gene 47 SEQ ID NO:57.
 XX
 XX Human; secreted protein; diagnosis; cystostatic; immunosuppressive;
 KW neotropic; neuroprotective; antiviral; antiallergic; hepatotropic;
 KW antidiabetic; antinflammatory; antitumor; antiparasitic; anticonvulsant;
 KW antibacterial; antifungal; antiparasitic; cardiant; gene therapy;
 KW food additive; preservative; chromosome identification; cancer;
 KW immune disorder; cardiovascular disorder; neurological disease;
 KW wound healing; infectious disease; ss.
 XX Homo sapiens.
 OS
 XX WO200058339-A2.
 PN
 XX 05-OCT-2000.
 PD
 XX 22-MAR-2000; 2000WO-US007440.
 PP
 XX 26-MAR-1999; 99US-0126503P.
 PR
 XX 17-DEC-1999; 99US-0172409P.
 PR
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX Rosen CA, Ruben SM, Komatsoulis G;
 PI
 XX WPI; 2000-594637/56.
 DR
 XX P-PSDB; AAB44642.
 DR
 XX Fifty nucleic acid molecules encoding human secreted proteins, useful in
 PT the prevention, treatment and diagnosis of cancer, immune disorders,
 PT cardiovascular disorders and neurological diseases.
 PT
 XX Claim 1; Page 364-365; 410pp; English.
 PS

XX The polynucleotide sequences given in AAC79681 to AAC79730 encode the
 CC human secreted proteins given in AAB44596 to AAB44645. AAB44646 to
 CC AAB44693 represent human secreted polypeptide sequences and proteins
 CC homologous to them, which are given in the exemplification of the present
 CC invention. Human secreted proteins have activities based on the tissues
 CC and cells the genes are expressed in. Examples of activities include:
 CC cytostatic; immunosuppressive; neotropic; neuroprotective; antiviral;
 CC antiallergic; hepatotropic; antidiabetic; antinflammatory; antitumor;
 CC anticonvulsant; antibacterial; antifungal; antiparasitic; and
 CC cardiant. The polynucleotides and polypeptides are useful for preventing,
 CC treating or ameliorating a medical condition in e.g. humans, mice,
 CC rabbits, goats, horses, cats, dogs, chickens or sheep. The polypeptides
 CC can also be used as a food additive or preservative to increase or
 CC decrease storage capabilities. The polynucleotides are useful for
 CC chromosome identification. They are also useful as probes for diagnosing
 CC a disorder related to the female reproductive system, particularly breast
 CC and/or ovarian cancer. They are also useful in the gene therapy of breast
 CC and ovarian cancer. The nucleic acids, protein, antibodies, agonists and
 CC antagonists from the present invention are useful in the diagnosis,
 CC treatment and prevention of: cancer; immune disorders; cardiovascular
 CC disorders; wound healing; neurological diseases; and infectious diseases.
 CC AAC79672 to AAC79680 and AAB44595 represent sequences used in the
 CC exemplification of the present invention
 XX
 XX Sequence 903 BP; 350 A; 162 C; 173 G; 218 T; 0 U; 0 Other;
 SQ
 Query Match 1.3%; Score 49; DB 3; Length 903;
 Best Local Similarity 100.0%; Pred. NO. 1.8e-07;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 QY 3714 AATGAAAAA 3762
 DB 788 AATGAAAAA 836
 RESULT 42
 AAH33675
 ID AAH33675 standard; cDNA; 1057 BP.
 XX
 XX AC AAH33675;
 XX
 XX 03-SEP-2001 (first entry)
 XX Human colon cancer antigen encoding cDNA SEQ ID NO:731.
 DE
 XX Human; colon cancer; colon cancer antigen; diagnosis; detection;
 KW colorectal carcinoma; ss.
 XX Homo sapiens.
 OS
 XX WO200122920-A2.
 PN
 XX 05-APR-2001.
 PD
 XX 28-SEP-2000; 2000WO-US026524.
 PP
 XX 29-SEP-1999; 99US-0157137P.
 PR
 XX 03-NOV-1999; 99US-0163280P.
 PR
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX Ruben SM, Barash SC, Birse CE, Rosen CA;
 PI
 XX WPI; 2001-235357/24.
 DR
 XX P-PSDB; AAG74244.
 DR
 XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
 PT useful for preventing, diagnosing and/or treating colorectal cancers.
 PT
 XX Claim 1; Page 2724; 9803pp; English.
 XX
 XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon

CC cancer-associated nucleic acid molecules (N) and proteins (P), where the
 CC proteins are collectively known as colon cancer antigens. The colon
 CC cancer antigens have cytostatic activity and can be used in gene therapy
 CC and vaccine production. N and P may be used in the prevention, diagnosis
 CC and treatment of diseases associated with inappropriate P expression. For
 CC example, N and P may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of P by expressing inactive proteins or to
 CC supplement the patient's own production of P. Additionally, N may be used
 CC to produce the colon cancer-associated Ps, by inserting the nucleic acids
 CC into a host cell and culturing the cell to express the proteins. N and P
 CC can be used in the prevention, diagnosis and treatment of colorectal
 CC carcinomas and cancers. AAH37196 to AAH37204 and AAH37789 represent
 CC sequences used in the exemplification of the present invention. N.B.
 CC Pages 666 to 682 and page 7053 of the sequence listing were missing at
 CC time of publication, meaning no sequences are present for SEQ ID NO:1027
 CC to 1052, 7921 and 7922

XX
 SQ Sequence 1057 BP; 318 A; 210 C; 207 G; 320 T; 0 U; 2 Other;

Query Match 1.3%; Score 49; DB 4; Length 1057;
 Best Local Similarity 100.0%; Pred. No. 1.9e-07;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA 3762
 DB 1006 AAATGAAAAA 1054

RESULT 43
 AAC79719
 ID AAC79719 standard; cDNA; 1263 BP.

XX AAC79719;

XX 12-FEB-2001 (first entry)

XX Human secreted protein gene 39 SEQ ID NO:49.

XX Human; secreted protein; diagnosis; cytostatic; immunosuppressive;
 KW neutropic; neuroprotective; antiviral; antiallergic; hepatotropic;
 KW antidiabetic; antiinflammatory; antitumor; anticonvulsant;
 KW antibacterial; antifungal; antiparasitic; cardiant; gene therapy;
 KW food additive; preservative; chromosome identification; cancer;
 KW immune disorder; cardiovascular disorder; neurological disease;
 KW wound healing; infectious disease; ss.

XX Homo sapiens.

XX WO200058339-A2.

XX 05-OCT-2000.

XX 22-MAR-2000; 2000WO-US007440.

XX 26-MAR-1999; 99US-0126503P.

XX 17-DEC-1999; 99US-0172409P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM, Komatsoulis G;

XX WPI; 2000-594637/56.

XX P-PSDB; AAB44634.

XX Fifty nucleic acid molecules encoding human secreted proteins, useful in
 PT the prevention, treatment and diagnosis of cancer, immune disorders,
 PT cardiovascular disorders and neurological diseases.

PS Claim 1; Page 360; 410pp; English.

XX The polynucleotide sequences given in AAC79681 to AAC79730 encode the
 CC human secreted proteins given in AAB44596 to AAB44645. AAB44646 to

CC AAB44693 represent human secreted polypeptide sequences and proteins
 CC homologous to them, which are given in the exemplification of the present
 CC invention. Human secreted proteins have activities based on the tissues
 CC and cells the genes are expressed in. Examples of activities include:
 CC cytostatic; immunosuppressive; neutropic; neuroprotective; antiviral;
 CC antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
 CC vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic; and
 CC cardiant. The polynucleotides and polypeptides are useful for preventing,
 CC treating or ameliorating a medical condition in e.g. humans, mice,
 CC rabbits, goats, horses, cats, dogs, chickens or sheep. The polypeptides
 CC can also be used as a food additive or preservative to increase or
 CC decrease storage capabilities. The polynucleotides are useful for
 CC chromosome identification. They are also useful as probes for diagnosing
 CC a disorder related to the female reproductive system, particularly breast
 CC and/or ovary cancer. They are also useful in the gene therapy of breast
 CC and ovarian cancer. The nucleic acids, protein, antibodies, agonists and
 CC antagonists from the present invention are useful in the diagnosis,
 CC treatment and prevention of: cancer; immune disorders; cardiovascular
 CC disorders; wound healing; neurological diseases; and infectious diseases.
 CC AAC79672 to AAC79680 and AAB44595 represent sequences used in the
 CC exemplification of the present invention

XX
 SQ Sequence 1263 BP; 306 A; 383 C; 323 G; 245 T; 0 U; 6 Other;

Query Match 1.3%; Score 49; DB 3; Length 1263;
 Best Local Similarity 100.0%; Pred. No. 1.7e-07;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA 3762
 DB 1189 AAATGAAAAA 1237

RESULT 44

AAC77616

ID AAC77616 standard; cDNA; 1373 BP.

XX AAC77616;

XX 08-FEB-2001 (first entry)

XX Human cancer associated gene sequence SEQ ID NO:10.

XX Human; cancer associated gene; cancer antigen; detection; cancer;
 KW diagnosis; cytostatic; proliferative; vulnary; immunomodulator;
 KW antidiabetic; antiasthmatic; antirheumatic; antiarthritic; antiviral;
 KW antiinflammatory; antithyroid; antiallergic; antibacterial; cardiant;
 KW dermatological; neuroprotective; thrombolytic; coagulant; neutropic;
 KW vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
 KW immune disorder; haematopoietic cell disorder; autoimmune disorder;
 KW allergic reaction; graft versus host disease; organ rejection;
 KW haemostatic; thrombolytic; cardiovascular disorder; infection;
 KW neurological disease; drug screening; ss.

XX Homo sapiens.

XX WO200055350-A1.

XX 21-SEP-2000.

XX 08-MAR-2000; 2000WO-US005882.

XX 12-MAR-1999; 99US-0124270P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM;

XX WPI; 2000-587533/55.

XX P-PSDB; AAB43407.

XX Novel isolated nucleic acids comprising sequences encoding peptides
 PT useful for treating or diagnosing e.g. cancer.

XX Claim 1; Page 616-617; 2352pp; English.

XX AAC77607 to AAC78448 encode the human cancer associated proteins given in

XX ABA43398 to ABA44239. The proteins can have activities based on the

XX tissues and cells the genes are expressed in. Example of activities

XX include: cytostatic; proliferative; vulnerary; immunomodulator;

XX antidiabetic; antiaesthetic; antirheumatic; antithyritic;

XX antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;

XX dermatological; neuroprotective; cardiant; thrombolytic; coagulant;

XX neotropic; vasotropic; antipsoriatic and antiangiogenic. The

XX polynucleotides and polypeptides can be used for preventing, treating or

XX ameliorating medical conditions and diagnosing pathological conditions.

XX Polynucleotides, polypeptides, antibodies, agonists and antagonists from

XX the present invention may be used to treat immune disorders by activating

XX or inhibiting the proliferation, differentiation or mobilisation of

XX immune cells, to treat disorders of haematopoietic cells, autoimmune

XX disorders, allergic reactions, graft versus host disease and organ

XX rejection, modulate haemostatic or thrombolytic activity, modulate

XX inflammation, cancers, cardiovascular disorders, neurological disease and

XX bacterial or viral infections. The peptides, nucleotides, antibodies,

XX agonists and antagonists may be also be used in drug screens. AAC78449 to

XX AAC78457 and ABA44240 represent sequences used in the exemplification of

XX the present invention

SQ Sequence 1373 BP; 398 A; 362 C; 370 G; 241 T; 0 U; 2 Other;

Query Match 1.3%; Score 49; DB 3; Length 1373;

Best Local Similarity 100.0%; Pred. No. 1.7e-07;

Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA 3762

DB 1294 AAATGAAAAA 1342

RESULT 45

ABV94043

ID ABV94043 standard; cDNA; 1429 BP.

XX AC ABV94043;

XX 08-JAN-2003 (first entry)

DE Breast carcinoma related nucleotide sequence SEQ ID NO:34.

DE Human; breast carcinoma; cancer; tumour; cytostatic; anti-tumour; gene;

XX ss.

XX Homo sapiens.

XX WO200246467-A2.

XX 13-JUN-2002.

XX 07-DEC-2001; 2001WO-IB002811.

XX 08-DEC-2000; 2000US-0254090P.

XX 07-DEC-2001; 2001US-00007926.

XX (IPSO-) IPSOGEN.

XX Bertucci F, Houlgatte R, Birnbaum D, Nguyen C, Viens P, Fert V;

XX WPI; 2002-619023/66.

XX Novel polynucleotide library useful in molecular characterization of a

XX carcinoma, comprising a pool of polynucleotide sequences or its

XX subences which are either underexpressed or overexpressed in tumor

XX cells.

XX Claim 1; Page 130-131; 401pp; English.

XX The present invention describes a polynucleotide library (I) useful in

XX the molecular characterisation of a carcinoma, comprising a pool of

XX polynucleotides or its subsequences which are either underexpressed or

XX overexpressed in tumour cells, and correspond to any of the

XX polynucleotide sequences chosen from the 468 sequences given in ABV94010

XX to ABV94477. Also described: (i) a polynucleotide array (II) useful for

XX the prognosis or diagnostic of tumour, comprising (I); and (2) detecting

XX (M1) differentially expressed polynucleotide sequences which are

XX correlated with a cancer, involves obtaining a polynucleotide sample from

XX a patient, and reacting the polynucleotide sample obtained with a probe

XX immobilised on a solid support, where the probe comprises any combination

XX of the polynucleotide sequences of (I) or its expression products encoded

XX by polynucleotide sequences of (I), and detecting the reaction product.

XX (i) have cytostatic activities and can be used as anti-tumour agents. (i)

XX is useful in molecular characterisation of a carcinoma. (I) and (II) are

XX useful for the prognosis or diagnostic of tumour, in differentiating a

XX normal cell from a cancer cell, detecting a hormone sensitive tumour

XX cell, differentiating a tumour with lymph nodes from a tumour without

XX lymph nodes, differentiating antitumour-sensitive tumours from

XX antitumour-insensitive tumours, and classifying good and poor prognosis

XX primary breast tumours. (i) is useful for large-scale molecular

XX characterisation of breast cancer that help in prediction, prognosis and

XX cancer treatment, and for detecting differentially expressed genes that

XX correlated with a cancer

SQ Sequence 1429 BP; 372 A; 384 C; 381 G; 292 T; 0 U; 0 Other;

Query Match 1.3%; Score 49; DB 6; Length 1429;

Best Local Similarity 100.0%; Pred. No. 1.7e-07;

Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA 3762

DB 1336 AAATGAAAAA 1384

RESULT 46

AAH72649

ID AAH72649 standard; cDNA; 1676 BP.

XX AC AAH72649;

XX 19-SEP-2001 (first entry)

XX Human cervical cancer marker nucleic acid 3923.

XX Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.

XX Homo sapiens.

XX WO200142467-A2.

XX 14-JUN-2001.

XX 08-DEC-2000; 2000WO-US033312.

XX 08-DEC-1999; 99US-0169681P.

XX 21-DEC-1999; 99US-0171350P.

XX 14-MAR-2000; 2000US-0189315P.

XX 12-MAY-2000; 2000US-0203791P.

XX 09-JUN-2000; 2000US-0210600P.

XX 21-JUL-2000; 2000US-0220114P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Schlegel R, Deeds J, Berger A, Zhao X;

XX WPI; 2001-375006/39.

XX New isolated nucleic acid for diagnosing and treating cervical cancer and

XX for assessing and detecting compounds for treating the cancer.

XX Claim 1; Page 780-781; 1051pp; English.

Db 1321 AATGAAAAA 1369

RESULT 49

ABV25281

ID ABV25281 standard; cDNA; 1702 BP.

XX AC ABV25281;

XX DT 16-SEP-2002 (first entry)

XX DE Human prostate expression marker cDNA 25272.

XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX KW pharmacogenomic marker; gene; ss.

XX OS Homo sapiens.

XX PN WO200160860-A2.

XX PD 23-AUG-2001.

XX PF 20-FEB-2001; 2001WO-US005171.

XX PR 17-FEB-2000; 2000US-0183319P.

XX PR 16-MAR-2000; 2000US-0189862P.

XX PR 25-MAY-2000; 2000US-0207454P.

XX PR 09-JUN-2000; 2000US-0211314P.

XX PR 18-JUL-2000; 2000US-0219007P.

XX PR 13-DEC-2000; 2000US-0255281P.

XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX PI Schlegel R, Endege WO, Monahan JE;

XX DR WPI; 2001-662795/76.

XX XX Novel isolated nucleic acid molecule associated with cancerous state of

XX PT prostate cells and correlating with presence of prostate cancer, useful

XX PT for detecting presence of prostate cancer, stage of prostate cancer.

XX XX Claim 1; Page 4967; 11750pp; English.

XX CC The invention relates to an isolated nucleic acid molecule (I) comprising

XX CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the

XX CC specification or its complement. (I) is useful for: (a) assessing whether

XX CC a patient is afflicted with prostate cancer; (b) monitoring the

XX CC progression of prostate cancer in a patient; (c) assessing the efficacy

XX CC of a test compound to inhibit prostate cancer in a patient; (d) assessing

XX CC the efficacy of a therapy for inhibiting prostate cancer in a patient;

XX CC (e) selecting a composition for inhibiting prostate cancer in a patient;

XX CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)

XX CC determining whether prostate cancer has metastasized in a patient; (h)

XX CC assessing the aggressiveness or indolence of prostate cancer in a patient

XX CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker

XX SQ Sequence 1702 BP; 560 A; 309 C; 393 G; 440 T; 0 U; 0 Other;

XX Query Match 1.3%; Score 49; DB 5; Length 1702;

XX Best Local Similarity 100.0%; Pred. No. 1.6e-07;

XX Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AATGAAAAA 3762

Db 1321 AATGAAAAA 1369

RESULT 50

ABV21893

ID ABV21893 standard; cDNA; 1702 BP.

XX AC ABV21893;

DT 13-SEP-2002 (first entry)

XX DE Human prostate expression marker cDNA 21894.

XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX KW pharmacogenomic marker; gene; ss.

XX OS Homo sapiens.

XX PN WO200160860-A2.

XX PD 23-AUG-2001.

XX PF 20-FEB-2001; 2001WO-US005171.

XX PR 17-FEB-2000; 2000US-0183319P.

XX PR 16-MAR-2000; 2000US-0189862P.

XX PR 25-MAY-2000; 2000US-0207454P.

XX PR 09-JUN-2000; 2000US-0211314P.

XX PR 18-JUL-2000; 2000US-0219007P.

XX PR 13-DEC-2000; 2000US-0255281P.

XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX PI Schlegel R, Endege WO, Monahan JE;

XX DR WPI; 2001-662795/76.

XX PT Novel isolated nucleic acid molecule associated with cancerous state of

XX PT prostate cells and correlating with presence of prostate cancer, useful

XX PT for detecting presence of prostate cancer, stage of prostate cancer.

XX XX Claim 1; Page 3725; 11750pp; English.

XX CC The invention relates to an isolated nucleic acid molecule (I) comprising

XX CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the

XX CC specification or its complement. (I) is useful for: (a) assessing whether

XX CC a patient is afflicted with prostate cancer; (b) monitoring the

XX CC progression of prostate cancer in a patient; (c) assessing the efficacy

XX CC of a test compound to inhibit prostate cancer in a patient; (d) assessing

XX CC the efficacy of a therapy for inhibiting prostate cancer in a patient;

XX CC (e) selecting a composition for inhibiting prostate cancer in a patient;

XX CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)

XX CC determining whether prostate cancer has metastasized in a patient; (h)

XX CC assessing the aggressiveness or indolence of prostate cancer in a patient

XX CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker

XX SQ Sequence 1702 BP; 560 A; 309 C; 393 G; 440 T; 0 U; 0 Other;

XX Query Match 1.3%; Score 49; DB 5; Length 1702;

XX Best Local Similarity 100.0%; Pred. No. 1.6e-07;

XX Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AATGAAAAA 3762

Db 1321 AATGAAAAA 1369

RESULT 51

AAD07369

ID AAD07369 standard; DNA; 1933 BP.

XX AC AAD07369;

XX DT 10-AUG-2001 (first entry)

XX DE Human DNA encoding SAP-1 protein (marker 20).

XX KW Human; cytostatic; diagnosis; prostate disorder; prostate cancer; BPH;
XX KW benign prostatic hyperplasia; SAP-1 protein; ds.

XX OS Homo sapiens.

```
PN WO200136674-A2.
XX
XX
XX 25-MAY-2001.
XX
XX
XX 08-NOV-2000; 2000WO-GB004267.
XX
XX 13-NOV-1999; 99GB-00026805.
XX
XX (ASTR ) ASTRAZENECA AB.
XX (ASTR ) ASTRAZENECA UK LTD.
XX
XX Bull JH, Ellison G, Paskins LD;
XX
XX WPI; 2001-343837/36.
XX
XX Diagnosing and monitoring prostate disorders, by analysis of 26 gene
XX transcripts that exhibit aberrant expression levels in prostate disorder
XX tissues, and provides a means of early diagnosis.
XX
XX Claim 1; Page 65; 69pp; English.
XX
XX The patent discloses a method for diagnosing, prognosing or monitoring a
XX prostate disorder which involves the analysis of 26 gene transcripts
XX (referred as markers) that exhibit aberrant expression levels in prostate
XX disorder tissues and provides a means of early diagnosis. This method is
XX useful for diagnosing, prognosing or monitoring a prostate disorder. It
XX also provides a means of distinguishing prostate cancer from benign
XX prostatic hyperplasia (BPH) and for identifying potential anti-prostate
XX disorder therapeutic compounds. The present sequence is a human DNA
XX encoding SAP-1 protein (referred as marker 20). The SAP-1 protein is
XX identified as a protein recruited by serum response factor to the c-fos
XX serum response element
XX
XX Sequence 1933 BP; 556 A; 486 C; 385 G; 506 T; 0 U; 0 Other;
XX
Query Match 1.3%; Score 49; DB 4; Length 1933;
Best Local Similarity 100.0%; Pred. No. 1.6e-07;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3714 AAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 1878 AAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1926
RESULT 52
ABK92253
ID ABK92253 standard; DNA; 1933 BP.
XX
XX AC ABK92253;
XX
XX DT 15-AUG-2002 (first entry)
XX
XX Prostate cancer-associated DNA sequence #139.
XX
XX Prostate cancer; prostate tumour tissue; human; mammal; cytostatic;
XX gene therapy; gene; ds.
XX
XX Mammalia.
XX
XX WO200230268-A2.
XX
XX 18-APR-2002.
XX
XX 12-OCT-2001; 2001WO-US032045.
XX
XX 13-OCT-2000; 2000US-00697576.
XX 08-DEC-2000; 2000US-00733288.
XX 08-DEC-2000; 2000US-00733742.
XX 24-JAN-2001; 2001US-02639572.
XX 16-MAR-2001; 2001US-0276791P.
XX 16-MAR-2001; 2001US-0276888P.
XX 06-APR-2001; 2001US-0281922P.
XX 24-APR-2001; 2001US-0286214P.
XX
PR 30-APR-2001; 2001US-00847046.
PR 04-MAY-2001; 2001US-0288589P.
XX
XX (EOGB-) EOS BIOTECHNOLOGY INC.
XX
XX Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;
XX
XX WPI; 2002-471335/50.
XX P-PSDB; ABG61934.
XX
XX Detecting a prostate cancer-associated transcript in a cell in a patient,
XX useful for diagnosing prostate cancer (PC) or screening modulators of PC,
XX by determining if prostate cancer-associated genes are expressed in a
XX prostate tissue.
XX
XX Claim 22; Page 416; 436pp; English.
XX
XX The present invention relates to methods of detecting a prostate cancer-
XX associated transcript in a cell from a patient. The method comprises
XX contacting a biological sample from the patient with prostate cancer-
XX associated polynucleotides (designated PC genes) that selectively
XX hybridise to a sequence that is at least 80% identical to them. The
XX prostate cancer-associated polynucleotide sequences are differentially
XX expressed in prostate tumour tissue or in prostate cancer and are derived
XX from the tissues of various organisms such as humans or other mammals
XX (e.g. mice, sheep and dogs). The methods of the invention are useful for
XX diagnosing and treating prostate cancer in mammals. The prostate cancer-
XX associated genes are useful for diagnosing or treating prostate cancer,
XX as well as for identifying modulators of prostate cancer or agents that
XX inhibit prostate cancer. The nucleic acid sequences are particularly
XX useful in gene therapy, as a vaccine or in antisense applications.
XX ABK92115-ABK92263 represent prostate cancer-associated polynucleotide
XX sequences
XX
XX Sequence 1933 BP; 556 A; 486 C; 385 G; 506 T; 0 U; 0 Other;
XX
Query Match 1.3%; Score 49; DB 5; Length 1933;
Best Local Similarity 100.0%; Pred. No. 1.6e-07;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3714 AAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 1878 AAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1926
RESULT 53
AAD16758
ID AAD16758 standard; cDNA; 1998 BP.
XX
XX AC AAD16758;
XX
XX DT 22-NOV-2001 (first entry)
XX
XX Human novel protein-encoding gene 9 cDNA clone HUCMO06, SEQ ID NO:19.
XX
XX Human; cytostatic; gene therapy; inflammatory disorder; neural disorder;
XX Parkinson's disease; Alzheimer's disease; immune system disorder; AIDS;
XX autoimmune disease; rheumatoid arthritis; muscular disorder; ischaemia;
XX reproductive disorder; Crohn's disease; pulmonary disorder; cancer;
XX myocardial infarction; glomerulonephritis; nephrotic syndrome; tumour;
XX haematopoietic disorder; rhinitis; asthma; diabetes; atherosclerosis;
XX anti-sense therapy; endocrine disorder; leukaemia; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 56..1174
XX /tag= a
XX /product= "Human novel protein"
XX /note= "CDS does not include start codon"
XX /partial
XX
XX WO200155202-A1.
```

```
XX PD 02-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US001325.
XX PF
XX PF
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205535P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239355P.
PR 13-OCT-2000; 2000US-023937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 08-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 11-DEC-2000; 2000US-0251990P.
PR 05-JAN-2001; 2001US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
PR (HUMA-) HUMAN GENOME SCI INC.
XX PA
XX PI
XX PI
XX PI
XX PI
```

DR WPI; 2001-451925/48.
DR P-PSDB; AAE09698.
XX Isolated polypeptide for treating, preventing and/ or prognosing medical
PT disorders and also for testing and detection e.g. diagnosis and screening
PT for agonists.
XX Claim 1; SEQ ID NO 19; 469pp; English.
XX
XX AAD16750-AAD16775 represent cDNAs corresponding to novel human protein
CC genes, and AAE09690-AA09715 represent the proteins they encode. AAD16777
CC -AAD16780 represent novel human genomic DNA fragments. The novel proteins
CC and their DNAs are useful for diagnosing, treating, preventing and/or
CC prognosing inflammatory disorders (bursitis or tendonitis); neural
CC disorders (e.g. Parkinson's disease, Alzheimer's disease); immune system
CC disorders; AIDS; autoimmune diseases (e.g., rheumatoid arthritis);
CC muscular disorders; reproductive disorders; gastrointestinal disorders
CC (malabsorption syndrome, Crohn's disease); pulmonary disorders;
CC cardiovascular disorders (myocardial infarction, ischaemia, arrhythmias);
CC renal disorders (glomerulonephritis, nephrotic syndrome); cancerous
CC disease and conditions (breast cancer); hyperproliferative disorders
CC (leukaemia, hyperplasia); tumours; foetal and developmental abnormalities
CC ; haematopoietic disorders; respiratory disorders (rhinitis, asthma);
CC angiotonic disorders; diabetes; atherosclerosis; endocrine disorders;
CC pregnancy-related disorders and infections. The novel protein DNA is
CC useful in gene therapy and anti-sense therapy. The proteins can also be
CC used to aid wound healing and epithelial cell proliferation, to prevent
CC skin aging due to sunburn, to maintain organs before transplantation, for

Query Match 1.3%; Score 49; DB 4; Length 1998;
Best Local Similarity 100.0%; Pred. No. 1.6e-07;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATCAA 3762
DB 1906 AAATCAA 1954

RESULT 54
ADC22050
ID ADC22050 standard; cDNA; 1998 BP.
AC ADC22050;
XX
XX
XX 18-DEC-2003 (first entry)
DE Human cDNA from secreted protein gene 9 #1.
XX Secreted protein; cytostatic; antibacterial; virucide; neuroprotective;
KW gynaecological; gastrointestinal-gen; cardiant; cardiovascular-gen;
KW nephrotropic; antiinflammatory; muscular-gen; respiratory-gen;
KW immunosuppressive; cerebroprotective; vasotropic; nootropic;
KW antiallergic; cancer; bacterial infection; viral infection;
KW neural disorder; immune system disorder; blood disorder;
KW muscular disorder; reproductive disorder; gastrointestinal disorder;
KW pulmonary disorder; cardiovascular disorder; renal disorder;
KW inflammatory disorder; proliferative disorder; Human; ss; gene.
XX
XX Homo sapiens.
XX
XX US2003082681-A1.
XX
XX 01-MAY-2003.
XX
XX 07-MAR-2002; 2002US-00091391.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 18-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227183P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-023497P.
PR 25-SEP-2000; 2000US-023498P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235634P.
PR 27-SEP-2000; 2000US-0235636P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 23-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.

PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241826P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249255P.
 PR 17-NOV-2000; 2000US-0249257P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 PR 17-JAN-2001; 2001US-00764903.
 PR (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM, Barash SC;
 XX
 DR WPI; 2003-786903/74.
 DR P-PSDB; ADC22076.
 XX
 PT New nucleic acid molecules and polypeptides for diagnosing, preventing or
 PT treating disorders associated with aberrant expression of the
 PT polypeptide, e.g. neural or cardiovascular disorders, and in chromosome
 PT identification.
 XX
 PS Claim 1; SEQ ID NO 19; 242bp; English.
 XX
 CC The invention relates to an isolated nucleic acid molecule (cDNA)

encoding a human secreted protein, representing one of 15 novel genes.
 Also included are recombinant vectors, host cells (expressing the
 protein), the secreted proteins (including their fragments, epitopes and
 homologues), an isolated antibody that binds specifically to the protein,
 diagnosing a pathological condition or susceptibility to a pathological
 condition (comprising determining the presence or absence of a mutation
 in the nucleic acid and diagnosing a condition based on the presence or
 absence of the mutation), diagnosing a pathological condition or
 susceptibility to a pathological condition (comprising determining the
 presence or amount of expression of the protein in a biological sample
 and diagnosing a condition based on the presence or amount of expression
 of the protein), preventing, treating or ameliorating a medical condition
 by administering the nucleic acid or protein to a mammalian subject,
 identifying a binding partner to the protein, the gene corresponding to
 the cDNA sequence, and identifying an activity in a biological assay
 (comprising expressing the nucleic acid in a cell, isolating the
 supernatant, detecting an activity in a biological assay and identifying
 the protein in the supernatant having the activity). The nucleic acids
 and proteins display the following activities: Cytostatic, antibacterial,
 CC Virucide, Neuroprotective, Gynaecological, Gastrointestinal-Gen,
 CC Cardiant, Cardiovascular-Gen, Nephrotropic, Antiinflammatory, Muscular-

Query Match 1.3%; Score 49; DB 9; Length 1998;
 Best Local Similarity 100.0%; Pred. No. 1.6e-07;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
 |||||
 DB 1906 AAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1954

RESULT 55
 ACID13422
 ID ACID13422 standard; cDNA; 2038 BP.
 XX
 AC ACID13422;
 XX
 DT 13-AUG-2003 (first entry)
 XX
 DE Human DNA encoding a p53 modifier, SEQ ID 94.
 XX
 KW Human; ss; gene; p53 modifier; cytostatic; cancer; cytostatic;
 KW antiangiogenic; antiapoptotic; p53 pathway; breast cancer; colon cancer;
 KW kidney cancer; lung cancer; ovarian cancer; angiogenesis; cell cycle;
 KW apoptotic disorder; cell proliferation disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200299122-A1.
 XX
 PD 12-DEC-2002.
 XX
 PF 03-JUN-2002; 2002WO-US017382.
 XX
 PR 05-JUN-2001; 2001US-0296076P.
 PR 10-OCT-2001; 2001US-0328605P.
 PR 15-FEB-2002; 2002US-0357253P.
 XX
 PA (EXEL-) EXELIXIS INC.
 XX
 PI Friedman L, Florman GD, Belvin M, Francis-Lang H, Li D, Funke RP;
 XX
 DR WPI; 2003-156859/15.
 DR P-PSDB; ABO07249.
 XX
 PT Identifying modulators of the p53 pathway for use in treating apoptotic
 PT or cell proliferation disorders, comprises screening for agents that
 PT modulate activity of a human ortholog of genes that modify the p53
 PT pathway in Drosophila.
 XX
 PS Example 2; Page 376-377; 678pp; English.
 XX
 CC The invention relates to identifying (M1) a candidate p53 pathway

modulating agent, by contacting an assay system comprising a purified HM polypeptide (human orthologue of genes that modify the p53 pathway in Drosophila) or nucleic acid with a test agent under conditions, where but for the presence of the test agent, the system provides a reference activity, and detecting a test-agent-biased activity of the assay system. Also included are modulating (M2) a p53 pathway of a cell (comprising contacting a cell defective in p53 function with a candidate modulator that specifically binds to a HM polypeptide comprising an HM amino acid sequence, where p53 function is restored), modulating (M3) a p53 pathway in a mammalian cell (comprising contacting the cell with an agent that specifically binds an HM polypeptide or nucleic acid) and diagnosing (M4) a disease in a patient (comprising: (a) obtaining a biological sample from the patient; (b) contacting the sample with a probe for HM expression; (c) comparing the results with a control; and (d) determining whether the comparison indicates a likelihood disease). (M1) is useful for identifying modulators of the p53 pathway. A probe for HM expression is useful for diagnosing breast, colon, kidney, lung and ovarian cancer, in a patient, where the cancer has greater than 25 % expression level. Modulators identified by (M1) are useful in a variety of diagnostic and therapeutic applications, where disease or disorder prognosis is related to defects in the p53 pathway, such as, angiogenesis, apoptotic or cell proliferation disorders (e.g. cancer). Another two new methods (M2 and M3) are useful for modulating the p53 pathway of a cell, thus restoring the p53 function of the cell, so that the cell undergoes normal proliferation or progression through the cell cycle. (M2) and (M3) are also useful for treating defects in the p53 pathway such as angiogenic, apoptotic or cell proliferation disorders. The present sequence is an HM nucleic acid encoding a p53 pathway modifying protein

Sequence 2038 BP; 527 A; 580 C; 567 G; 364 T; 0 U; 0 Other;

Query Match 1.3%; Score 49; DB 7; Length 2038;
Best Local Similarity 100.0%; Pred. No. 1.6e-07;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA 3762
DB 1974 AAATGAAAAA 2022

RESULT 56

ABZ71965
ID ABZ71965 standard; cDNA; 2106 BP.

AC ABZ71965;

DT 01-APR-2003 (first entry)

XX Human cDNA for transketolase GenBank X67688.

XX Human; cancer; stomach cancer; cytostatic; gene; ss.

XX Homo sapiens.

XX WO200283899-A1.

XX 24-OCT-2002.

XX 28-MAR-2002; 2002WO-JP003038.

XX 10-APR-2001; 2001JP-00112039.

XX 21-SEP-2001; 2001JP-00280193.

XX (TAKA-) TAKARA BIO INC.

XX Yoshikawa Y, Okamoto S, Oura T, Mineno J, Asada K, Kato I;

XX Inoue H, Mori M;

XX WPI; 2003-093022/08.

XX Measuring changes in expression of 264 cancer associated genes for
PT detection of stomach cancer and screening of potential anticancer agents.

PS Claim 2; Page; 266pp; Japanese.

XX The invention relates to a method for the detection of cancer in which a
CC change in the expression of 1 or more of 264 specified cancer associated
CC genes, ABZ71694-ABZ71957, or of sequences at least 80% homologous to them
CC in the specimen tissue as compared to normal tissue is observed. The
CC genes are used in detection, diagnosis and treatment of cancer,
CC especially of stomach cancer. The present sequence is that of a cancer
CC associated polynucleotide of the invention. Note: The present sequence
CC was not given in the printed specification but was isolated using the
CC GenBank accession number given in the DE line

XX Sequence 2106 BP; 529 A; 608 C; 589 G; 380 T; 0 U; 0 Other;

Query Match 1.3%; Score 49; DB 7; Length 2106;
Best Local Similarity 100.0%; Pred. No. 1.6e-07;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA 3762
DB 2044 AAATGAAAAA 2092

RESULT 57

ABQ54860/c

ID ABQ54860 standard; cDNA; 2143 BP.

XX AC ABQ54860;

XX 22-AUG-2002 (first entry)

XX Human ovarian antigen HAZAA31 cDNA, SEQ ID NO:740.

XX Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
XX ovarian cancer; breast cancer; tumour; reproductive system disorder;
XX infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
XX PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;
XX inflammatory condition; immune disorder; blood disorder;
XX cardiovascular disorder; respiratory disorder; neurological disorder;
XX gastrointestinal disorder; urinary system disorder; drug screening;
XX gene therapy; chromosome mapping; forensic analysis;
XX antibody preparation; cytostatic; immunomodulatory; neuroprotective;
XX antiinflammatory; gynaecological; reproductive; chromosome 3p14.3; gene;
XX ss.

XX Homo sapiens.

XX WO200200677-A1.

XX 03-JAN-2002.

XX 07-JUN-2001; 2001WO-US018569.

XX 07-JUN-2000; 2000US-0209467P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Birse CE, Rosen CA;

XX WPI; 2002-147878/19.

XX P-PSDB; ABP41783.

XX Isolated nucleic acid molecules encoding novel ovarian polypeptides,
XX useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian
XX cancer), immune disorders, cardiovascular disorders and neurological
XX diseases.

XX Claim 1; SEQ ID NO 740; 2922pp; English.

XX The invention relates to 2175 novel human ovarian antigens (ABP41054-
XX ABP43228) and to cDNAs encoding them (ABQ5431-ABQ56305), and also
XX encompasses polypeptides 90% identical and polynucleotides 95% identical
XX to the sequences of the invention. The invention additionally relates to

CC recombinant vectors and host cells comprising human ovarian antigen
 CC polynucleotides, antibodies against human ovarian antigens, and the use
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
 CC treating, prognosing or preventing various ovary and/or breast-related
 CC disorders. Such conditions include ovarian cancer and breast cancer, and
 CC metastatic tumours of ovarian or breast origin, reproductive system
 CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
 CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
 CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
 CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
 CC vaginitis), immune disorders (e.g., congenital and acquired
 CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
 CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
 CC respiratory disorders, neurological disorders, gastrointestinal disorders
 CC and urinary system disorders. Ovarian antigen polypeptides and
 CC polynucleotides may also be used in screening for compounds which
 CC modulate ovarian antigen expression or activity. The polynucleotides may
 CC further be used for gene therapy, chromosome mapping, in the
 CC identification of individuals and in forensic analysis, and the
 CC polypeptides may be used as food additives or to prepare antibodies
 CC useful in disease diagnosis, drug targeting and phenotyping. The present
 CC sequence represents cDNA encoding a human ovarian antigen of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 2143 BP; 377 A; 598 C; 620 G; 545 T; 0 U; 3 Other;

Query Match 1.3%; Score 49; DB 6; Length 2143;
 Best Local Similarity 100.0%; Pred. No. 1.6e-07;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AATGAAAAA 3762

DB 80 AATGAAAAA 32

RESULT 58

AAC76644

ID AAC76644 standard; cDNA; 2543 BP.

XX AC AAC76644;

XX DT 08-FEB-2001 (first entry)

XX DE Human ORFX ORF2199 polynucleotide sequence SEQ ID NO:4397.

XX Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
 KW vulnery; antipsoriatic; antiparkinsonian; nontropic; neuroprotective;
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
 KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
 KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
 KW thrombosis; contraceptive; ss.

XX Homo sapiens.

XX WO200058473-A2.

XX PN 05-OCT-2000.

XX PD 31-MAR-2000; 2000WO-US008621.

XX PF 31-MAR-1999; 99US-0127607P.

XX PR 02-APR-1999; 99US-0127636P.

XX PR 05-APR-1999; 99US-0127728P.

PR 30-MAR-2000; 2000US-00540763.

XX (CURA-) CURAGEN CORP.

XX PI Shimkets RA, Leach M;

XX WPI; 2000-602362/57.

XX P-PSDB; AAB42435.

XX Novel nucleic acids and peptides derived from open reading frame X,

XX useful for treating e.g. cancers, proliferative disorders,

XX neurodegenerative disorders and cardiovascular disease.

XX Claim 5; Page 3587-3589; 5507pp; English.

XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,

XX which represent the human ORFX open reading frames 1 to 3161. The ORFX

XX sequences have activities such as: cytostatic; hepatotropic; vulnery;

XX antipsoriatic; antiparkinsonian; nontropic; neuroprotective; osteopathic;

XX anticonvulsant; antiarthritic; immunosuppressant; immunostimulant;

XX cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;

XX dermatological; immunosuppressive; antiinflammatory; antibacterial;

XX antiviral; antifungal; antirheumatic; antithyroid; and antianemic. The

XX sequences can be used for determining the presence of or predisposition

XX to, or preventing or treating pathological conditions associated with an

XX ORFX-associated disorder. The nucleic acids can be used to express ORFX

XX proteins in gene therapy vectors. The proteins and nucleic acids may be

XX used to treat cancers, proliferative disorders, neurodegenerative

XX disorders, osteoarthritis, graft vs host disease, cardiovascular disease,

XX diabetes mellitus, hypertension, hypothyroidism, cholesterol ester

XX storage, systemic lupus erythematosus, severe combined immunodeficiency

XX (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune

XX disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and

XX cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to

XX enhance coagulation; to inhibit thrombosis; and as a contraceptive

XX SQ Sequence 2543 BP; 630 A; 589 C; 684 G; 639 T; 0 U; 1 Other;

Query Match 1.3%; Score 49; DB 3; Length 2543;

Best Local Similarity 100.0%; Pred. No. 1.5e-07;

Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AATGAAAAA 3762

DB 2494 AATGAAAAA 2542

RESULT 59

AAD43556

ID AAD43556 standard; cDNA; 2713 BP.

XX AC AAD43556;

XX DT 14-NOV-2002 (first entry)

XX DE Human CD2000 cDNA.

XX Human; immunoglobulin; Ig; SLAM associated protein; SAP; CD2000 protein;

XX immune proliferative disorder; immune disorder; rheumatoid arthritis;

XX carcinoma; autoimmune disorder; multiple sclerosis; Grave's disease;

XX Hashimoto's disease; acquired immune deficiency syndrome; hepatotropic;

XX osteoarthritis; allergic inflammatory disorder; viral infection; asthma;

XX psoriasis; apoptotic disorder; systemic lupus erythematosus; bronchitis;

XX diabetes mellitus; septic shock; chronic obstructive pulmonary disease;

XX emphysema; cachexia; hepatic circulatory disorder; hepatitis; cirrhosis;

XX acute myeloid leukaemia; haemophilia; anaemia; gene therapy; cytostatic;

XX immunosuppressive; neuroprotective; antiinflammatory; Crohn's disease;

XX osteopathic; antibacterial; immunomodulator; inflammatory bowel disease;

XX jaundice; dermatological; ulcerative colitis; AIDS; chromosome 1q23;

XX Gene; ss.

XX OS Homo sapiens.

AAD43565
 ID AAD43565 standard; DNA; 2713 BP.
 AC AAD43565;
 XX
 DT 14-NOV-2002 (first entry)
 XX
 DE Human CD2000 DNA #2.
 XX
 KW Human; immunoglobulin; Ig; SLAM associated protein; SAP; CD2000 protein; immune proliferative disorder; immune disorder; rheumatoid arthritis; carcinoma; autoimmune disorder; multiple sclerosis; Grave's disease; Hashimoto's disease; acquired immune deficiency syndrome; hepatotropic; osteoarthritis; allergic inflammatory disorder; viral infection; asthma; psoriasis; apoptotic disorder; systemic lupus erythematosus; bronchitis; diabetes mellitus; septic shock; chronic obstructive pulmonary disease; emphysema; cachexia; hepatic circulatory disorder; hepatitis; cirrhosis; acute myeloid leukaemia; haemophilia; anaemia; gene therapy; cytostatic; immunosuppressive; neuroprotective; antiinflammatory; Crohn's disease; osteopathic; antibacterial; immunomodulator; inflammatory bowel disease; jaundice; dermatological; ulcerative colitis; AIDS; gens; ds.
 XX
 OS Homo sapiens.
 XX
 PN EP1223218-A1.
 XX
 PD 17-JUL-2002.
 XX
 PF 02-NOV-2001; 2001EP-00309339.
 XX
 PR 03-NOV-2000; 2000US-00706167.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Fraser CC;
 XX
 WPI; 2002-620680/67.
 XX
 Novel isolated polypeptide containing immunoglobulin and immunoglobulin-like domains and SLAM associated protein, termed CD2000 or CD2001, useful for treating immune, inflammatory, or hepatic circulatory disorders.
 XX
 PS Disclosure; Page 78; 138pp; English.
 XX
 CC The invention relates to nucleic acid molecule, designated CD2000 which encodes a polypeptide containing immunoglobulin (Ig) and Ig-like domains and SLAM associated protein (SAP) motifs. CD2000 DNA and protein is useful for treating disorder such as immune proliferative disorders, immune disorders (e.g. carcinoma), viral infection, autoimmune disorders (e.g., arthritis, multiple sclerosis, Grave's disease, and Hashimoto's disease), T cell disorder (e.g. acquired immune deficiency syndrome (AIDS)), inflammatory bowel disease (e.g. Crohn's disease and ulcerative colitis), inflammatory disorders (e.g. rheumatoid arthritis and psoriasis), apoptotic disorders (e.g. systemic lupus erythematosus, and osteoarthritis), allergic inflammatory disorders (e.g. asthma and emphysema), insulin-dependent diabetes mellitus, cytotoxic disorders, septic shock, chronic obstructive pulmonary disease (e.g. emphysema), bronchitis, cachexia, jaundice, hepatic circulatory disorders, hepatitis, cirrhosis, acute myeloid leukaemia, haemophilia and anaemia. CD2000 DNA is used in gene therapy. CD2000 DNA is useful in screening assays, detection assays (e.g. chromosomal mapping, tissue typing, forensic biology), predictive medicine (e.g. diagnostic assays, prognostic assays, monitoring clinical trials and pharmacogenomics), and in methods of treatment (e.g. therapeutic and prophylactic). The present sequence is human CD2000 DNA
 XX
 SQ Sequence 2713 BP; 798 A; 648 C; 529 G; 738 T; 0 U; 0 Other;
 Query Match 1.3%; Score 49; DB 6; Length 2713;
 Best Local Similarity 100.0%; Pred. No. 1.5e-07;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AATCGAAAAA
 (|||||)

Db 2650 AATCGAAAAA
 RESULT 62
 AAD43564
 ID AAD43564 standard; DNA; 2713 BP.
 XX
 AC AAD43564;
 XX
 DT 14-NOV-2002 (first entry)
 XX
 DE Human CD2000 DNA #1.
 XX
 KW Human; immunoglobulin; Ig; SLAM associated protein; SAP; CD2000 protein; immune proliferative disorder; immune disorder; rheumatoid arthritis; carcinoma; autoimmune disorder; multiple sclerosis; Grave's disease; Hashimoto's disease; acquired immune deficiency syndrome; hepatotropic; osteoarthritis; allergic inflammatory disorder; viral infection; asthma; psoriasis; apoptotic disorder; systemic lupus erythematosus; bronchitis; diabetes mellitus; septic shock; chronic obstructive pulmonary disease; emphysema; cachexia; hepatic circulatory disorder; hepatitis; cirrhosis; acute myeloid leukaemia; haemophilia; anaemia; gene therapy; cytostatic; immunosuppressive; neuroprotective; antiinflammatory; Crohn's disease; osteopathic; antibacterial; immunomodulator; inflammatory bowel disease; jaundice; dermatological; ulcerative colitis; AIDS; gens; ds.
 XX
 OS Homo sapiens.
 XX
 PN EP1223218-A1.
 XX
 PD 17-JUL-2002.
 XX
 PF 02-NOV-2001; 2001EP-00309339.
 XX
 PR 03-NOV-2000; 2000US-00706167.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Fraser CC;
 XX
 WPI; 2002-620680/67.
 XX
 Novel isolated polypeptide containing immunoglobulin and immunoglobulin-like domains and SLAM associated protein, termed CD2000 or CD2001, useful for treating immune, inflammatory, or hepatic circulatory disorders.
 XX
 PS Disclosure; Page 77-78; 138pp; English.
 XX
 CC The invention relates to nucleic acid molecule, designated CD2000 which encodes a polypeptide containing immunoglobulin (Ig) and Ig-like domains and SLAM associated protein (SAP) motifs. CD2000 DNA and protein is useful for treating disorder such as immune proliferative disorders, immune disorders (e.g. carcinoma), viral infection, autoimmune disorders (e.g., arthritis, multiple sclerosis, Grave's disease, and Hashimoto's disease), T cell disorder (e.g. acquired immune deficiency syndrome (AIDS)), inflammatory bowel disease (e.g. Crohn's disease and ulcerative colitis), inflammatory disorders (e.g. rheumatoid arthritis and psoriasis), apoptotic disorders (e.g. systemic lupus erythematosus, and insulin-dependent diabetes mellitus), cytotoxic disorders, septic shock, chronic obstructive pulmonary disease (e.g. emphysema), bronchitis, cachexia, jaundice, hepatic circulatory disorders, hepatitis, cirrhosis, acute myeloid leukaemia, haemophilia and anaemia. CD2000 DNA is used in gene therapy. CD2000 DNA is useful in screening assays, detection assays (e.g. chromosomal mapping, tissue typing, forensic biology), predictive medicine (e.g. diagnostic assays, prognostic assays, monitoring clinical trials and pharmacogenomics), and in methods of treatment (e.g. therapeutic and prophylactic). The present sequence is human CD2000 DNA
 XX
 SQ Sequence 2713 BP; 799 A; 646 C; 529 G; 739 T; 0 U; 0 Other;
 Query Match 1.3%; Score 49; DB 6; Length 2713;
 Best Local Similarity 100.0%; Pred. No. 1.5e-07;

Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA 1.3%; Score 49; DB 6; Length 2713;
Best Local Similarity 100.0%; Pred. No. 1.5e-07;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 2650 AAATGAAAAA 1.3%; Score 49; DB 6; Length 2713;
Best Local Similarity 100.0%; Pred. No. 1.5e-07;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 63

AAAD3567
ID AAD43567 standard; DNA; 2713 BP.

XX AC AAD43567;

XX DT 14-NOV-2002 (first entry)

XX DE Human CD2000 DNA #4.

XX Human; immunoglobulin; Ig; SLAM associated protein; SAP; CD2000 protein;
immune proliferative disorder; immune disorder; rheumatoid arthritis;
carcinoma; autoimmune disorder; multiple sclerosis; Grave's disease;
Hashimoto's disease; acquired immune deficiency syndrome; hepatotropic;
osteoarthritis; allergic inflammatory disorder; viral infection; asthma;
psoriasis; apoptotic disorder; systemic lupus erythematosus; bronchitis;
diabetes mellitus; septic shock; chronic obstructive pulmonary disease;
emphysema; cachexia; hepatic circulatory disorder; hepatitis; cirrhosis;
acute myeloid leukaemia; haemophilia; anaemia; gene therapy; cytostatic;
immunopressive; neutropenic; antineoplastic; Crohn's disease;
osteopathic; antibacterial; immunomodulator; inflammatory bowel disease;
jaundice; dermatological; ulcerative colitis; AIDS; gene; ds.

XX OS Homo sapiens.

XX FN EP123218-A1.

XX PD 17-JUL-2002.

XX PF 02-NOV-2001; 2001EP-00309339.

XX PR 03-NOV-2000; 2000US-00706167.

XX PA (MILL-) MILLENNIUM PHARM INC.

XX PI Fraser CC;

XX DR WPI; 2002-620680/67.

XX Novel isolated polypeptide containing immunoglobulin and immunoglobulin-
like domains and SLAM associated protein, termed CD2000 or CD2001, useful
for treating immune, inflammatory, or hepatic circulatory disorders.

XX PS Disclosure; Page 79-80; 139pp; English.

XX The invention relates to nucleic acid molecule, designated CD2000 which
encodes a polypeptide containing immunoglobulin (Ig) and Ig-like domains
and SLAM associated protein (SAP) motifs. CD2000 DNA and protein is
useful for treating disorder such as immune proliferative disorders,
immune disorders (e.g. carcinoma), viral infection, autoimmune disorders
(e.g., arthritis, multiple sclerosis, Grave's disease, and Hashimoto's
disease), T cell disorder (e.g. acquired immune deficiency syndrome
(AIDS)), inflammatory bowel disease (e.g. Crohn's disease and ulcerative
colitis), inflammatory disorders (e.g. rheumatoid arthritis and
psoriasis), apoptotic disorders (e.g. systemic lupus erythematosus, and
osteoarthritis), allergic inflammatory disorders (e.g. emphysema), bronchitis,
insulin-dependent diabetes mellitus (e.g. emphysema), hepatitis, cirrhosis,
chronic obstructive pulmonary disease (e.g. emphysema), hepatitis, cirrhosis,
cachexia, jaundice, hepatic circulatory disorders, hepatitis, cirrhosis,
acute myeloid leukaemia, haemophilia and anaemia. CD2000 DNA is used in
gene therapy. CD2000 DNA is useful in screening assays, detection assays
(e.g. chromosomal mapping, tissue typing, forensic biology), predictive
medicine (e.g. diagnostic assays, prognostic assays, monitoring clinical
trials and pharmacogenomics), and in methods of treatment (e.g.
therapeutic and prophylactic). The present sequence is human CD2000 DNA

SQ Sequence 2713 BP; 799 A; 648 C; 529 G; 737 T; 0 U; 0 Other;

Query Match 1.3%; Score 49; DB 6; Length 2713;
Best Local Similarity 100.0%; Pred. No. 1.5e-07;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA 1.3%; Score 49; DB 6; Length 2713;
Best Local Similarity 100.0%; Pred. No. 1.5e-07;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 2650 AAATGAAAAA 1.3%; Score 49; DB 6; Length 2713;
Best Local Similarity 100.0%; Pred. No. 1.5e-07;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 64

ABN86559
ID ABN86559 standard; DNA; 2851 BP.

XX AC ABN86559;

XX DT 05-NOV-2002 (first entry)

XX Canine low affinity IgE receptor (CD23) nCaCD23_2851 DNA sequence.

XX Canine; immunoglobulin E; IgE; CD23; antiallergic; antiaesthetic; gene;
antiinflammatory; dermatological; gene therapy; vaccine; ds.

XX Canis familiaris.

XX Key Location/Qualifiers

XX CDS 199..1077

XX FT /*tag= a

XX FT /product= "CD23 receptor"

XX US6410714-B1.

XX 25-JUN-2002.

XX 24-MAR-2000; 2000US-00535521.

XX 24-MAR-1999; 99US-0125913P.

XX (HESK-) HESKA CORP.

XX Weber ER, McCall CA;

XX WPI; 2002-588996/53.

XX P-PSDB; ABB81056.

XX New isolated canine low affinity immunoglobulin E receptor nucleic acid
molecule, useful for protecting canids from diseases mediated by the
receptor, such as allergy, atopic dermatitis, asthma, and hay fever.

XX Claim 1a; Col 31-36; 33pp; English.

XX The invention relates to isolated canine low affinity immunoglobulin E
(IgE) receptor (CD23) polypeptides and encoding nucleic acid molecules.
The CD23 polypeptides can be expressed by standard recombinant
methodology. The CD23 polynucleotides are useful for protecting canids
from diseases mediated by CD23, for developing compounds that regulate
IgE and/or CD23 levels in a canid for treating allergy related diseases
such as atopic dermatitis, asthma, hay fever and food sensitivities. The
present sequence represents a CD23 nucleic acid molecule nCaCD23_2851

XX SQ Sequence 2851 BP; 604 A; 938 C; 781 G; 528 T; 0 U; 0 Other;

Query Match 1.3%; Score 49; DB 6; Length 2851;
Best Local Similarity 100.0%; Pred. No. 1.5e-07;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA 1.3%; Score 49; DB 6; Length 2851;
Best Local Similarity 100.0%; Pred. No. 1.5e-07;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 2802 AAATGAAAAA 1.3%; Score 49; DB 6; Length 2851;
Best Local Similarity 100.0%; Pred. No. 1.5e-07;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 65

ABN86560/c
ID ABN86560 standard; DNA; 2851 BP.

XX AC ABN86560;

XX DT 05-NOV-2002 (first entry)

XX DE Canine CD23 nCaCD23_2851 complementary DNA sequence.

XX KW Canine; immunoglobulin E; IgE; CD23; antiallergic; antiasthmatic;
XX XW antiinflammatory; dermatological; gene therapy; vaccine; ds.

XX OS Canis familiaris.

XX FN US6410714-B1.

XX PD 25-JUN-2002.

XX PF 24-MAR-2000; 2000US-00535521.

XX FR 24-MAR-1999; 99US-0125913P.

XX FA (HESK-) HESKA CORP.

XX PI Weber ER, McCall CA;

XX DR WPI; 2002-588896/63.

XX PT New isolated canine low affinity immunoglobulin E receptor nucleic acid
XX FT molecule, useful for protecting canids from diseases mediated by the
XX PT receptor, such as allergy, atopic dermatitis, asthma, and hay fever.

XX PS Claim 1b; Col 37-40; 33pp; English.

XX CC The invention relates to isolated canine low affinity immunoglobulin E
XX CC (IgE) receptor (CD23) polypeptides and encoding nucleic acid molecules.
XX CC The CD23 polypeptides can be expressed by standard recombinant
XX CC methodology. The CD23 polynucleotides are useful for protecting canids
XX CC from diseases mediated by CD23, for developing compounds that regulate
XX CC IgE and/or CD23 levels in a canid for treating allergy related diseases
XX CC such as atopic dermatitis, asthma, hay fever and food sensitivities. The
XX CC present sequence represents the complementary sequence of the CD23
XX CC nucleic acid molecule nCaCD23_2851

XX SQ Sequence 2851 BP; 528 A; 781 C; 938 G; 604 T; 0 U; 0 Other;

Query Match 1.3%; Score 49; DB 6; Length 2851;

Best Local Similarity 100.0%; Pred. No. 1.5e-07;

Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3714 AAATGAAAAA 3762

Db 50 AAATGAAAAA 2

RESULT 66

ABT18794/c

ID ABT18794 standard; DNA; 2944 BP.

XX AC ABT18794;

XX DT 16-APR-2003 (first entry)

XX DE Aspergillus fumigatus essential gene #1152.

XX XW Fungicide; cytostatic; essential gene; Aspergillus fumigatus; infection;
XX KW cancer; contamination; biofilm; antibody; immune response; ds.

XX OS Aspergillus fumigatus.

XX FN WO200286090-A2.

XX PD 31-OCT-2002.

XX PF

XX 23-APR-2002; 2002WO-US013142.

XX PR 23-APR-2001; 2001US-0285697P.

XX PR 27-APR-2001; 2001US-0287066P.

XX PR 05-JUN-2001; 2001US-0295890P.

XX PR 09-JUL-2001; 2001US-0303899P.

XX PR 31-AUG-2001; 2001US-0316362P.

XX PA (ELIT-) ELITRA PHARM INC.

XX PI Jiang B, Tishkoff D, Zamudio C, Eroshkin AM, Hu W, Lemieux SM;

XX DR WPI; 2003-093124/08.

XX PT New purified or isolated nucleic acids of essential genes of Aspergillus

XX PT fumigatus, useful for treating or preventing infections by A. fumigatus,

XX PT or for treating a non-infectious disease in a subject e.g. cancer.

XX PS Disclosure; Page; 175pp; English.

XX CC The invention relates to novel purified or isolated nucleic acids of

XX CC essential genes of Aspergillus fumigatus. The isolated nucleic acids of

XX CC the invention are used to treat or prevent infections by a pathogenic

XX CC organism such as A. fumigatus, to treat a non-infectious disease in a

XX CC subject (e.g. cancer), to prevent or contain contamination of an object

XX CC by A. fumigatus, or to prevent or inhibit formation on a surface of a

XX CC biofilm comprising A. fumigatus. The polynucleotides are useful for

XX CC expressing recombinant protein for characterization, screening or

XX CC therapeutic use, as markers for host tissues in which the pathogenic

XX CC organisms invade or reside, for comparing with the DNA sequence of A.

XX CC fumigatus to identify duplicated genes or paralogues having the same or

XX CC similar biochemical activity and/or function, for comparing with DNA

XX CC sequences of other related or distant pathogenic organisms to identify

XX CC potential orthologous essential or virulence genes, for selecting and

XX CC making oligomers for attachment to a nucleic acid array for examination

XX CC of expression patterns, for raising anti-protein antibodies, as an

XX CC antigen to raise anti-DNA antibodies or to elicit another immune

XX CC response, and for identifying polynucleotides encoding the other protein

XX CC with which binding occurs or to identify inhibitors of the binding

XX CC interaction. The polypeptides may be used to raise antibodies or to

XX CC elicit immune responses, as a reagent in assays designed to quantitatively

XX CC determine levels of the protein in biological fluids, as a marker for

XX CC host tissues in which pathogenic organism invade or reside, and to

XX CC isolate correlative receptors or ligands in the case of virulence

XX CC factors. This polynucleotide sequence represents one of the essential

XX CC genes of Aspergillus fumigatus of the invention

XX SQ Sequence 2944 BP; 676 A; 770 C; 765 G; 733 T; 0 U; 0 Other;

Query Match 1.3%; Score 49; DB 7; Length 2944;

Best Local Similarity 100.0%; Pred. No. 1.5e-07;

Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3714 AAATGAAAAA 3762

Db 941 AAATGAAAAA 893

RESULT 67

ABT20610/c

ID ABT20610 standard; DNA; 3362 BP.

XX AC ABT20610;

XX DT 16-APR-2003 (first entry)

XX DE Aspergillus fumigatus essential gene #2968.

XX XW Fungicide; cytostatic; essential gene; Aspergillus fumigatus; infection;
XX KW cancer; contamination; biofilm; antibody; immune response; ds.

XX OS Aspergillus fumigatus.

XX WO200286090-A2.
XX 31-OCT-2002.
XX 23-APR-2002; 2002WO-US013142.
XX 23-APR-2001; 2001US-0285697P.
XX 27-APR-2001; 2001US-0287066P.
XX 05-JUN-2001; 2001US-0295890P.
XX 09-JUL-2001; 2001US-0303893P.
XX 31-AUG-2001; 2001US-0316362P.
XX (ELIT-) ELITRA PHARM INC.
XX Jiang B, Tishkoff D, Zamudio C, Eroshkin AM, Hu W, Lemieux SM;
XX WPI; 2003-093124/08.
XX New purified or isolated nucleic acids of essential genes of *Aspergillus*
XX *fumigatus*, useful for treating or preventing infections by *A. fumigatus*,
XX or for treating a non-infectious disease in a subject e.g. cancer.
XX Disclosure; Page; 175pp; English.
XX The invention relates to novel purified or isolated nucleic acids of
XX essential genes of *Aspergillus fumigatus*. The isolated nucleic acids of
XX the invention are used to treat or prevent infections by a pathogenic
XX organism such as *A. fumigatus*, to treat a non-infectious disease in a
XX subject (e.g. cancer), to prevent or contain contamination of an object
XX by *A. fumigatus*, or to prevent or inhibit formation on a surface of a
XX biofilm comprising *A. fumigatus*. The polynucleotides are useful for
XX expressing recombinant protein for characterisation, screening or
XX therapeutic use, as markers for host tissues in which the pathogenic
XX organisms invade or reside, for comparing with the DNA sequence of *A.*
XX *fumigatus* to identify duplicated genes or paralogues having the same or
XX similar biochemical activity and/or function, for comparing with DNA
XX sequences of other related or distant pathogenic organisms to identify
XX potential orthologous essential or virulence genes, for selecting and
XX making oligomers for attachment to a nucleic acid array for examination
XX of expression patterns, for raising anti-protein antibodies, as an
XX antigen to raise anti-DNA antibodies or to elicit another immune
XX response, and for identifying polynucleotides encoding the other protein
XX with which binding occurs or to identify inhibitors of the binding
XX interaction. The polypeptides may be used to raise antibodies or to
XX elicit immune response, as a reagent in assays designed to quantitatively
XX determine levels of the protein in biological fluids, as a marker for
XX host tissues in which pathogenic organisms invade or reside, and to
XX isolate correlative receptors or ligands in the case of virulence
XX factors. This polynucleotide sequence represents one of the essential
XX genes of *Aspergillus fumigatus* of the invention
XX Sequence 3362 BP; 774 A; 901 C; 867 G; 820 T; 0 U; 0 Other;
Query Match 1.3%; Score 49; DB 7; Length 3362;
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3714 AAATGAAAAA
DB 1359 AAATGAAAAA
RESULT 68
ABQ92014
ID ABQ92014 standard; cDNA; 4236 BP.
XX AC ABQ92014;
XX 04-OCT-2002 (first entry)
XX Human polynucleotide SEQ ID NO 11.

KW Human; cytostatic; antirheumatic; antiarthritic; vulnary; analgesic;
KW antiinflammatory; antibacterial; immunosuppressive; antiparkinsonian;
KW neuroprotective; nootropic; osteopathic; haemostatic; vasotropic;
KW antiulcer; fungicide; antidiabetic; antiaesthetic; antiallergic;
KW immunostimulant; antiparasitic; secreted protein; transmembrane protein;
KW cytokine; cell proliferation; cell differentiation; autoimmune disease;
KW stem cell; growth factor; nervous system disease; neuropathy;
KW Alzheimer's disease; Parkinson's disease; Huntington's disease;
KW osteoporosis; severe combined immunodeficiency; SCID; infection;
KW multiple sclerosis; rheumatoid arthritis; gene therapy; gene; ss.
XX Homo sapiens.
XX US2002065394-A1.
XX 30-MAY-2002.
XX 22-DEC-2000; 2000US-00745763.
XX 18-MAR-1998; 98US-00040963.
XX (JACO/) JACOBS K.
XX (MCCO/) MCCOY J M.
XX (LAVA/) LAVALAJE E R.
XX (COLL/) COLLINS-RACIE L A.
XX (EVAN/) EVANS C.
XX (MERE/) MERBERG D.
XX (TREA/) TREACY M.
XX (SPAU/) SPAULDING V.
XX Jacobs K, Mccoy JM, Lavallie ER, Collins-Racie LA, Evans C;
XX Merberg D, Treacy M, Spaulding V;
XX WPI; 2002-582343/62.
XX P-PSDB; ABPe1797.
XX Novel secreted or transmembrane protein and polynucleotide encoding the
XX protein, useful for diagnosis and treatment of neurological disorders,
XX cancer, autoimmune diseases, bone disorders and lung or liver fibrosis.
XX Claim 34; Page 106-108; 284pp; English.
XX The invention relates to human secreted or transmembrane protein (I),
XX their fragments and is encoded by specific complementary deoxyribonucleic
XX acid (cDNA) inserts (II), where the protein is substantially free from
XX other mammalian proteins. (I) are useful for preventing, treating or
XX ameliorating a medical condition, especially immunological treatment or
XX prevention of tumours. (I) exhibits activity relating to angiogenesis,
XX cytokine, cell proliferation, cell differentiation, antiinflammatory,
XX stem cell growth factor activity and activin or inhibin-related
XX activities. (I) can be used to manipulate stem cells in culture to give
XX rise to neuroepithelial cells that can be used to augment or replace
XX cells damaged by illness, autoimmune disease, accidental damage or
XX genetic disorders. (I) induces the proliferation of neural cells and
XX regeneration of nerve and brain tissue and is useful for the treatment of
XX central and peripheral nervous system diseases and neuropathies, such as
XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX lateral sclerosis. (I) is involved in chemotactic or chemokinetic
XX activity, regulation of haematopoiesis and is useful for treating myeloid
XX or lymphoid cell disorders, platelet disorders such as thrombocytopaenia
XX and for regeneration of bone, cartilage, tendon, ligament and/or nerve
XX tissue growth and in tissue repair, healing of burns, incisions, ulcers,
XX for treating osteoporosis, osteoarthritis, bone degenerative disorders or
XX periodontal disease. (I) is also useful for gut protection or
XX regeneration and treatment of lung or liver fibrosis, reperfusion injury
XX in various tissues, various immune deficiencies and disorders including
XX severe combined immunodeficiency (SCID), bacterial or fungal infections,
XX autoimmune disorders e.g. multiple sclerosis, rheumatoid arthritis,
XX diabetes mellitus, myasthenia gravis, allergic reactions and conditions,
XX such as asthma or other respiratory problems (II) is useful to express
XX recombinant protein, as markers for tissues in which the corresponding
XX protein is preferentially expressed and in gene therapy. The present
XX sequence is that of a polynucleotide of the invention

CC host tissues in which pathogenic organism invade or reside, and to
CC isolate correlative receptors or ligands in the case of virulence
CC factors. This polynucleotide sequence represents one of the essential
CC genes of Aspergillus fumigatus of the invention

XX
SQ Sequence 4944 BP; 1177 A; 1248 C; 1244 G; 1275 T; 0 U; 0 Other;
Query Match 1.3%; Score 49; DB 7; Length 4944;
Best Local Similarity 100.0%; Pred. No. 1.3e-07;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps

OY 3714 AAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
| | | | |
Db 1941 AAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1893

RESULT 70
ABT20014/C
ID ABT20014 standard; DNA; 5362 BP.
XX AC ABT20014;
XX DT 16-APR-2003 (first entry)
DE Aspergillus fumigatus essential gene #2372.
XX Fungicide; cytostatic; essential gene; Aspergillus fumigatus; infection;
KW cancer; contamination; biofilm; antibody; immune response; db.
XX OS Aspergillus fumigatus.
XX WO200286090-A2.
XX PD 31-OCT-2002.
XX PF 23-APR-2002; 2002WO-US013142.
XX PR 23-APR-2001; 2001US-0285697P.
XX PR 27-APR-2001; 2001US-0287066P.
XX PR 05-JUN-2001; 2001US-0295890P.
XX PR 09-JUL-2001; 2001US-0303899P.
XX PR 31-AUG-2001; 2001US-0316362P.
XX PA (ELIT-) ELITRA PHARM INC.
XX PI Jiang B, Tishkoff D, Zamudio C, Eroshkin AM, Hu W, Lemieux SM;
XX WIPI; 2003-093124/08.

XX New purified or isolated nucleic acids of essential genes of Aspergillus
XX fumigatus, useful for treating or preventing infections by A. fumigatus,
XX or for treating a non-infectious disease in a subject e.g. cancer.

XX Disclosure; Page; 175pp; English.

XX The invention relates to novel purified or isolated nucleic acids of
XX essential genes of Aspergillus fumigatus. The isolated nucleic acids of
XX the invention are used to treat or prevent infections by a pathogenic
XX organism such as A. fumigatus, to treat a non-infectious disease in a
XX subject (e.g. cancer), to prevent or contain contamination of an object
XX by A. fumigatus, or to prevent or inhibit formation on a surface of a
XX biofilm comprising A. fumigatus. The polynucleotides are useful for
XX expressing recombinant protein for characterisation, screening or
XX therapeutic use, as markers for host tissues in which the pathogenic
XX organisms invade or reside, for comparing with the DNA sequence of A.
XX fumigatus to identify duplicated genes or paralogues having the same or
XX similar biochemical activity and/or function, for comparing with DNA
XX sequences of other related or distant pathogenic organisms to identify
XX potential orthologous essential or virulence genes, for selecting and
XX making oligomers for attachment to a nucleic acid array for examination
XX of expression patterns, for raising anti-protein antibodies, as an
XX antigen to raise anti-DNA antibodies or to elicit another immune
XX response, and for identifying polynucleotides encoding the other proteins

CC with which binding occurs or to identify inhibitors of the binding
CC interaction. The polypeptides may be used to raise antibodies or to
CC elicit immune response, as a reagent in assays designed to quantitatively
CC determine levels of the protein in biological fluids, as a marker for
CC host tissues in which pathogenic organism invade or reside, and to
CC isolate correlative receptors or ligands in the case of virulence
CC factors. This polynucleotide sequence represents one of the essential
CC genes of *Aspergillus fumigatus* of the invention
XX
SQ Sequence 5362 BP; 1321 A; 1343 C; 1334 G; 1364 T; 0 U; 0 Other;

Query Match 1.3%; Score 49; DB 7; Length 5362;
Best Local Similarity 100.0%; Pred. No. 1.3e-07;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AATGCAA 3762
DB 2359 AATGCAA 2311

RESULT 71
AAH69364
ID AAH69364 standard; cDNA; 295 BP.
XX
AC AAH69364;
XX
DT 19-SEP-2001 (first entry)
XX
DE Human cervical cancer marker nucleic acid 638.
XX
KW Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN WO200142467-A2.
XX
PD 14-JUN-2001.
XX
PF 08-DEC-2000; 2000WO-US033312.
XX
PR 08-DEC-1999; 99US-0169681P.
PR 21-DEC-1999; 99US-0171350P.
PR 14-MAR-2000; 2000US-0169315P.
PR 12-MAY-2000; 2000US-0203791P.
PR 09-JUN-2000; 2000US-0210600P.
PR 21-JUL-2000; 2000US-0220114P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Deeds J, Berger A, Zhao X;
XX
DR WPI; 2001-375006/39.

XX New isolated nucleic acid for diagnosing and treating cervical cancer and
XX for assessing and detecting compounds for treating the cancer.

XX Claim 1; Page 215; 1051pp; English.

XX The invention relates to novel genes (AAH68727-AAH73383) associated with
XX cervical cancer with cytostatic activity. The nucleic acids and encoded
XX polypeptides are useful: to assess if a patient is afflicted with
XX cervical cancer or has a pre-malignant condition; to monitor the
XX progression of cervical cancer or a premalignant condition in a patient;
XX and to select and/or assess the efficacy of a compound or therapy for
XX inhibiting cervical cancer in a patient. The nucleic acids may also be
XX useful for gene therapy

XX Sequence 295 BP; 133 A; 46 G; 47 C; 69 T; 0 U; 0 Other;
Query Match 1.3%; Score 48; DB 4; Length 295;
Best Local Similarity 100.0%; Pred. No. 5.1e-07;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGCAA 3762
DB 233 AATGCAA 280

RESULT 72
ABX40871
ID ABX40871 standard; cDNA; 337 BP.
XX
AC ABX40871;
XX
DT 20-FEB-2003 (first entry)
XX
DE Bovine EST associated with lactation/muscle/fat deposition #6036.

XX Bovine; ss; EST; expressed sequence tag; lactation; LMFD;
KW muscle deposition; fat deposition; genome mapping; gene identification;
KW gene analysis; cattle breeding.
XX
OS Bos Taurus.

XX US2002137139-A1.
XX
XX 26-SEP-2002.
XX
XX 24-SEP-2001; 2001US-00960352.
XX
XX 12-JAN-1999; 99US-0115707P.
PR
PR 11-JAN-2000; 2000US-00480902.

XX (BYAT/) BYATT J C.
XX (MATH/) MATHIALAGAN N.
XX (TAON/) TAO N.
XX (WARE/) WARREN W C.

XX Byatt JC, Mathialagan N, Tao N, Warren WC;
XX
XX WPI; 2003-110599/10.

XX New nucleic acid associated with lactation, and muscle and fat
XX deposition, useful for genome mapping, gene identification and analysis,
XX cattle breeding, or for genetically improving cattle.

XX Claim 2; SEQ ID NO 6036; 245pp; English.

XX The invention relates to a purified nucleic acid molecule associated with
XX lactation or muscle and fat deposition (designated LMFD), derived from
XX cattle, and the LMFD nucleic acid can specifically hybridize to a second
XX nucleic acid molecule comprising any of 15112 nucleotide sequences,
XX appearing as ABX34836-ABX49947, or complements of them. Also included are
XX ; (1) a transformed cell having a nucleic acid comprising an LMFD nucleic
XX acid linked to a promoter and a 3' non-translated sequence that
XX functions in the cell to cause termination of transcription and addition
XX of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
XX (2) determining a level or pattern of a molecule in a bovine cell or
XX tissue comprising: (a) incubating a marker nucleic acid (comprising any
XX of the 15112 nucleic acid sequences or its complement or fragment) with a
XX complementary nucleic acid molecule obtained from the bovine cell or
XX tissue, where hybridisation between the marker nucleic acid and the
XX complementary nucleic acid permits the detection of the molecule; and (b)
XX detecting the level or pattern of the complementary nucleic acid, where
XX the detection of the complementary nucleic acid is predictive of the
XX level or pattern of the molecule. The LMFD nucleic acid is used for
XX determining a level or pattern of a molecule in a bovine cell or tissue.
XX It is useful for genome mapping, gene identification and analysis, cattle
XX breeding, preparation of constructs for use in cattle gene expression, or
XX for genetically improving cattle. The present sequence is one of the
XX 15112 bovine LMFD EST (expressed sequence tag) nucleic acids. Note: The
XX present sequence was not shown in the USPTO web site:
XX seqdata.uspto.gov/sequence.html?DocID=20020137139

XX Sequence 337 BP; 150 A; 35 C; 40 G; 112 T; 0 U; 0 Other;

ID AAI87343 standard: cDNA; 418 BP.

```
XX AC AAT87343;
XX OS
XX PN 06-NOV-2001 (first entry)
XX DE Human polynucleotide SEQ ID NO 7403.
XX KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX KW tissue growth factor; immunomodulatory; cancer; leukaemia;
XX KW nervous system disorders; arthritis; inflammation; ss.
XX OS Homo sapiens.
XX PN WO200164835-A2.
XX PD 07-SEP-2001.
XX PF 26-FEB-2001; 2001WO-US004927.
XX PR 28-FEB-2000; 2000US-00515126.
XX PR 18-MAY-2000; 2000US-00577409.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Liu C, Drmanac RT;
XX DR WPI; 2001-514838/56.
XX DR P-PSDB; AAO07412.
XX PT Isolated nucleic acids and polypeptides, useful for preventing diagnosing
XX PT and treating e.g. leukemia, inflammation and immune disorders.
XX PS Claim 1; SEQ ID NO 7403; 1399pp + Sequence Listing; English.
XX CC The invention relates to human polynucleotides (AAI79941-AAI93841) and
XX CC the encoded proteins (AAO0010-AAO13910) that exhibit activity relating to
XX CC cytokine, cell proliferation or cell differentiation or which may induce
XX CC production of other cytokines in other cell populations. The
XX CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX CC peptide therapy. The polypeptides have various cytokine-like activities,
XX CC e.g. stem cell growth factor activity, haematopoiesis regulating
XX CC activity, tissue growth factor activity, immunomodulatory activity and/or
XX CC activin/inhibin activity and may be useful in the diagnosis and/or
XX CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX CC inflammation. Note: The sequence data for this patent did not form part
XX CC of the printed specification, but was obtained in electronic format
XX CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 418 BP; 222 A; 43 C; 62 G; 91 T; 0 U; 0 Other;

Query Match 1.3%; Score 48; DB 4; Length 418;
Best Local Similarity 100.0%; Pred. No. 4.8e-07;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
DB 272 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 319

RESULT 76
AAI88543
ID AAI88543 standard; cDNA; 419 BP.
XX AC AAI88543;
XX PN 06-NOV-2001 (first entry)
XX DE Human polynucleotide SEQ ID NO 8603.
XX KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX KW tissue growth factor; immunomodulatory; cancer; leukaemia;
```

```
KW nervous system disorders; arthritis; inflammation; ss.
XX OS Homo sapiens.
XX PN WO200164835-A2.
XX PD 07-SEP-2001.
XX PF 26-FEB-2001; 2001WO-US004927.
XX PR 28-FEB-2000; 2000US-00515126.
XX PR 18-MAY-2000; 2000US-00577409.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Liu C, Drmanac RT;
XX DR WPI; 2001-514838/56.
XX DR P-PSDB; AAO08612.
XX PT Isolated nucleic acids and polypeptides, useful for preventing diagnosing
XX PT and treating e.g. leukemia, inflammation and immune disorders.
XX PS Claim 1; SEQ ID NO 8603; 1399pp + Sequence Listing; English.
XX CC The invention relates to human polynucleotides (AAI79941-AAI93841) and
XX CC the encoded proteins (AAO0010-AAO13910) that exhibit activity relating to
XX CC cytokine, cell proliferation or cell differentiation or which may induce
XX CC production of other cytokines in other cell populations. The
XX CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX CC peptide therapy. The polypeptides have various cytokine-like activities,
XX CC e.g. stem cell growth factor activity, haematopoiesis regulating
XX CC activity, tissue growth factor activity, immunomodulatory activity and/or
XX CC activin/inhibin activity and may be useful in the diagnosis and/or
XX CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX CC inflammation. Note: The sequence data for this patent did not form part
XX CC of the printed specification, but was obtained in electronic format
XX CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 419 BP; 174 A; 78 C; 86 G; 91 T; 0 U; 0 Other;

Query Match 1.3%; Score 48; DB 4; Length 419;
Best Local Similarity 100.0%; Pred. No. 4.8e-07;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
DB 219 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 266

RESULT 77
AAH71167
ID AAH71167 standard; cDNA; 546 BP.
XX AC AAH71167;
XX PN 19-SEP-2001 (first entry)
XX DE Human cervical cancer marker nucleic acid 2441.
XX KW Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
XX OS Homo sapiens.
XX PN WO200142467-A2.
XX PD 14-JUN-2001.
XX PF 08-DEC-2000; 2000WO-US033312.
XX PR 08-DEC-1999; 99US-0169681P.
XX PR 21-DEC-1999; 99US-0171350P.
XX PR 14-MAR-2000; 2000US-0189315P.
```


(YUY/) YU Y.
PA (NAME/) RAMEKA J G.
PA (PAGE/) PAGE A.
PA (MATH/) MATHW A V.
PA (LEDF/) LEDFORD B L.
PA (WOES/) WOESSNER J P.
PA (HAAS/) HAAS W D.
PA (GARC/) GARCIA C A.
PA (KRICK/) KRICKER M.
PA (SLAT/) SLATER T.
PA (DAVI/) DAVIS K R.
PA (ALLE/) ALLEN K.
PA (HOFF/) HOFFMAN N.
PA (HURE/) HUREAN P.

Gorlach J, An Y, Hamilton CM, Price JL, Raines TM, Yu Y;
Rameka JG, Page A, Mathew AV, Ledford BL, Woessner JP, Haas WD;
Garcia CA, Kricker M, Slater T, Davis KR, Allen K, Hoffman N,
Hurban P;
WPI; 2002-403163/43.

New Arabidopsis thaliana nucleic acid for identifying homologous genes,
producing compositions that modulate the expression or function of its
encoded protein, and mapping functional regions of protein.

Claim 1; SEQ ID NO 579; 49pp + Sequence Listing; English.

The invention relates to an Arabidopsis thaliana nucleic acid (I)
comprising a sequence capable of hybridising under stringent conditions
to a sequence selected from any one of 999 sequences (ABN98233-ABN99231),
given in the specification or its fragment. A polypeptide (II) encoded by
(I) a transgenic plant (III) comprising an exogenous nucleic acid or a
genetically modified cell (IV) comprising an exogenous nucleic acid, is
useful for screening a candidate agent for its biological effect. (I) is
useful in identifying homologous or related genes, in producing
compositions that modulate the expression or function of its encoded
protein, mapping functional regions of the protein and in studying
associated physiological pathways. (I) is also useful for the genetic
manipulation of cells, particularly plant cells. (I) is also useful in
screening assays of various plant strains to determine the strains that
are best capable of withstanding a particular disease or environmental
stress. (II) and (III) are useful for screening of biologically active
agents, e.g. fungicides, insecticides, etc., for elucidating biochemical
pathways. The screened agents are useful in improved methods of treating
crops to prevent or treat disease. (II) are also useful in screening
programs to identify agents that mimic or enhance the action of tolerance
factors. Such agents are useful in improved methods of treating crops to
enhance their tolerance to environmental stress. (I) is also useful for
enhancing or inhibiting production of a biosynthetic product in a plant.
(III) is useful for identifying other mediators that may induce in a
expression of proteins of interest, for establishing the extent to which
any specific insect and/or pathogen is responsible for damage to a
particular plant, for identifying other mediators that enhance or induce
tolerance to environmental stress, for identifying factors involved in
biosynthetic pathways of nutritional, commercial, or medicinal value and
for identifying productions of nutritional, commercial or medicinal
value. (IV) is useful in the study of genetic function and regulation,
for alteration of the cellular metabolism and for screening compounds
that may affect the biological function of the gene or gene products.
Note: The sequence data for this patent did not form part of the printed
specification, but was obtained in electronic format directly from USPTO
at seqdata.uspto.gov/sequence.html?DocID=999909770445

Sequence 865 BP; 231 A; 191 C; 152 G; 291 T; 0 U; 0 Other;

Query Match 1.3%; Score 48; DB 6; Length 865;
Best Local Similarity 100.0%; Pred. No. 4.le-07;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAAGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
DB 54 AATGAAGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 7

RESULT 80
AAI87708
ID AAI87708 standard; cDNA; 903 BP.
XX
AC AAI87708;
DT 06-NOV-2001 (first entry)
XX Human polynucleotide SEQ ID NO 7768.
DE
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation; ss.
OS Homo sapiens.
XX WO200164835-A2.
PN
XX 07-SEP-2001.
PD
XX 26-FEB-2001; 2001WO-US004927.
XX PF
XX 28-FEB-2000; 2000US-00515126.
PR
XX 18-MAY-2000; 2000US-00577409.
XX (HYSE-) HYSEQ INC.
PA Tang YT, Liu C, Drmanac RT;
XX WPI; 2001-514838/56.
DR P-PSDB; AA007777.
XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing
PT and treating e.g. leukemia, inflammation and immune disorders.
XX Claim 1; SEQ ID NO 7768; 1399pp + Sequence Listing; English.
XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to
CC cytokines, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The vaccines or
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and/or
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at fcp.wipo.int/pub/published_pct_sequences
XX Sequence 903 BP; 297 A; 208 C; 192 G; 199 T; 0 U; 7 Other;
SQ Query Match 1.3%; Score 48; DB 4; Length 903;
Best Local Similarity 100.0%; Pred. No. 4.le-07;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAAGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
DB 220 AATGAAGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 267

RESULT 81
AAA93117
ID AAA93117 standard; cDNA; 1136 BP.
XX
AC AAA93117;
DT 12-JAN-2001 (first entry)
XX

(YUY/) YU Y.
PA (NAME/) RAMEKA J G.
PA (PAGE/) PAGE A.
PA (MATH/) MATHW A V.
PA (LEDF/) LEDFORD B L.
PA (WOES/) WOESSNER J P.
PA (HAAS/) HAAS W D.
PA (GARC/) GARCIA C A.
PA (KRICK/) KRICKER M.
PA (SLAT/) SLATER T.
PA (DAVI/) DAVIS K R.
PA (ALLE/) ALLEN K.
PA (HOFF/) HOFFMAN N.
PA (HURE/) HUREAN P.

Gorlach J, An Y, Hamilton CM, Price JL, Raines TM, Yu Y;
Rameka JG, Page A, Mathew AV, Ledford BL, Woessner JP, Haas WD;
Garcia CA, Kricker M, Slater T, Davis KR, Allen K, Hoffman N,
Hurban P;
WPI; 2002-403163/43.

New Arabidopsis thaliana nucleic acid for identifying homologous genes,
producing compositions that modulate the expression or function of its
encoded protein, and mapping functional regions of protein.

Claim 1; SEQ ID NO 579; 49pp + Sequence Listing; English.

The invention relates to an Arabidopsis thaliana nucleic acid (I)
comprising a sequence capable of hybridising under stringent conditions
to a sequence selected from any one of 999 sequences (ABN98233-ABN99231),
given in the specification or its fragment. A polypeptide (II) encoded by
(I) a transgenic plant (III) comprising an exogenous nucleic acid or a
genetically modified cell (IV) comprising an exogenous nucleic acid, is
useful for screening a candidate agent for its biological effect. (I) is
useful in identifying homologous or related genes, in producing
compositions that modulate the expression or function of its encoded
protein, mapping functional regions of the protein and in studying
associated physiological pathways. (I) is also useful for the genetic
manipulation of cells, particularly plant cells. (I) is also useful in
screening assays of various plant strains to determine the strains that
are best capable of withstanding a particular disease or environmental
stress. (II) and (III) are useful for screening of biologically active
agents, e.g. fungicides, insecticides, etc., for elucidating biochemical
pathways. The screened agents are useful in improved methods of treating
crops to prevent or treat disease. (II) are also useful in screening
programs to identify agents that mimic or enhance the action of tolerance
factors. Such agents are useful in improved methods of treating crops to
enhance their tolerance to environmental stress. (I) is also useful for
enhancing or inhibiting production of a biosynthetic product in a plant.
(III) is useful for identifying other mediators that may induce in a
expression of proteins of interest, for establishing the extent to which
any specific insect and/or pathogen is responsible for damage to a
particular plant, for identifying other mediators that enhance or induce
tolerance to environmental stress, for identifying factors involved in
biosynthetic pathways of nutritional, commercial, or medicinal value and
for identifying productions of nutritional, commercial or medicinal
value. (IV) is useful in the study of genetic function and regulation,
for alteration of the cellular metabolism and for screening compounds
that may affect the biological function of the gene or gene products.
Note: The sequence data for this patent did not form part of the printed
specification, but was obtained in electronic format directly from USPTO
at seqdata.uspto.gov/sequence.html?DocID=999909770445

Sequence 865 BP; 231 A; 191 C; 152 G; 291 T; 0 U; 0 Other;

Query Match 1.3%; Score 48; DB 6; Length 865;
Best Local Similarity 100.0%; Pred. No. 4.le-07;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAAA 3762
DB 54 AATGAAA 7

RESULT 80
AAI87708
ID AAI87708 standard; cDNA; 903 BP.
XX
AC AAI87708;
DT 06-NOV-2001 (first entry)
XX Human polynucleotide SEQ ID NO 7768.
DE
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation; ss.
XX Homo sapiens.
OS
XX WO200164835-A2.
PN
XX 07-SEP-2001.
PD
XX 26-FEB-2001; 2001WO-US004927.
XX PF
XX 28-FEB-2000; 2000US-00515126.
PR
XX 18-MAY-2000; 2000US-00577409.
PR
XX (HYSE-) HYSEQ INC.
PA
XX Tang YT, Liu C, Drmanac RT;
XX WPI; 2001-514838/56.
DR
P-PSDB; AA007777.
XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing
PT and treating e.g. leukemia, inflammation and immune disorders.
XX
XX Claim 1; SEQ ID NO 7768; 1399pp + Sequence Listing; English.
XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to
CC cytokines, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The vaccines or
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at fcp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 903 BP; 297 A; 208 C; 192 G; 199 T; 0 U; 7 Other;
SQ
Query Match 1.3%; Score 48; DB 4; Length 903;
Best Local Similarity 100.0%; Pred. No. 4.le-07;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAAA 3762
DB 220 AATGAAA 267

RESULT 81
AAA93117
ID AAA93117 standard; cDNA; 1136 BP.
XX
AC AAA93117;
DT 12-JAN-2001 (first entry)
XX


```
DE Human secreted protein coding sequence SEQ ID NO: 33.
XX
XX Human; secreted protein; cytokine; cell proliferation;
KW nutritional supplement; immune modulation; autoimmune disorder;
KW haematopoiesis regulation; tissue growth; haemostasis; inflammation; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 93..266
XX /tag= a
XX /product= "secreted protein"
XX sig_peptide 135..173
XX /tag= b
XX mat_peptide 174..263
XX /tag= c
XX
XX WO200049134-A1.
XX
XX 24-AUG-2000.
XX
XX 18-FEB-2000; 2000WO-US004340.
XX
XX 19-FEB-1999; 99US-0120680P.
XX 23-APR-1999; 99US-00398733.
XX 17-AUG-1999; 99US-0149639P.
XX 23-SEP-1999; 99US-0155686P.
XX 01-OCT-1999; 99US-0157247P.
XX 29-NOV-1999; 99US-0167822P.
XX 29-NOV-1999; 99US-0167823P.
XX 15-FEB-2000; 2000US-0182711P.
XX
XX (ALPH-) ALPHAGENE INC.
XX
XX Valenzuela D, Yuan O, Hoffman H, Hall J, Rapijko P;
XX WPI; 2000-549267/50.
XX P-PSDB; AAB23617.
XX
XX New secreted proteins and polynucleotides encoding them, which are
XX derived from Homosapiens, useful for therapy, diagnosis, and research, as
XX well as nutritional sources or supplements.
XX
XX Claim 42; Page 264; 309pp; English.
XX
XX The present sequence is the coding sequence for a human secreted protein.
XX The sequence was isolated from an adult lung cDNA library. The proteins
XX and coding sequences of the invention can be used in the isolation of
XX similar genes and proteins, in the elucidation of their function in vivo,
XX and to treat a number of conditions. It is possible that they may have
XX uses as nutritional supplements, as cytokine or cell proliferation
XX factors, in immune modulation, where they may be used to treat immune and
XX autoimmune diseases, as haematopoiesis regulators (treating myeloid or
XX lymphoid cell deficiencies), in the promotion of tissue growth, they may
XX have chemokine or chemotactic activity, haemostatic or thrombolytic
XX activity, or anti-inflammatory activity
XX
XX Sequence 1136 BP; 351 A; 253 C; 249 G; 283 T; 0 U; 0 Other;
SQ
Query Match 1.3%; Score 48; DB 3; Length 1136;
Best Local Similarity 100.0%; Pred. No. 3.9e-07;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3715 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 1080 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1127
RESULT 82
AAC79042
ID AAC79042 standard; DNA; 1467 BP.
XX
XX AAC79042;
XX
XX 14-FEB-2001 (first entry)
XX
XX Human secreted protein gene 46 clone HPWA089.
XX
XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
XX anti-allergic; hepatotropic; antidiabetic; anti-inflammatory; anti-ulcer;
XX vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;
XX cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
XX neurological disease; infection; human; secreted protein; ss.
XX
XX Homo sapiens.
XX
XX WO200058358-A1.
XX
XX 05-OCT-2000.
XX
XX 23-MAR-2000; 2000WO-US007725.
XX
XX 26-MAR-1999; 99US-0126602P.
XX 14-JAN-2000; 2000US-0176063P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM, Komatsoulis G;
XX WPI; 2000-594640/56.
XX P-PSDB; AAB44379.
XX
XX Fourty nine nucleic acid molecules encoding human secreted proteins,
XX useful in the prevention, treatment and diagnosis of cancer, immune
XX disorders, cardiovascular disorders and neurological diseases.
XX
XX Claim 1; Page 336-337; 367pp; English.
XX
XX The invention relates to the isolation of genes AAC78997-C79045 encoding
XX 49 human secreted proteins AAB4335-B44382. The genes can be used to
XX generate fusion proteins by linking to the gene for the human
XX immunoglobulin G Fc portion (AAC78988) for increasing the stability of
XX the fusion protein as compared to the human protein only. The genes and
XX proteins are useful for preventing, ameliorating or treating medical
XX conditions, e.g. by protein or gene therapy. The genes are isolated from
XX a range of human tissues disclosed in the specification. The nucleic
XX acids, proteins, antibodies and (ant)agonists are useful in the
XX diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
XX ovarian cancer, and other cancers of the adrenal gland, bone, bone
XX marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b)
XX immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic
XX anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
XX multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
XX cardiovascular disorders such as myocardial ischaemias; (d) wound healing
XX ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
XX infectious diseases such as viral, bacterial, fungal and parasitic
XX infections
XX
XX Sequence 1467 BP; 429 A; 288 C; 294 G; 447 T; 0 U; 9 Other;
SQ
Query Match 1.3%; Score 48; DB 3; Length 1467;
Best Local Similarity 100.0%; Pred. No. 3.7e-07;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3715 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 1414 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1461
RESULT 83
ABK87557
ID ABK87557 standard; cDNA; 1510 BP.
XX
XX ABK87557;
XX
XX 24-SEP-2002 (first entry)
XX
```

```

XX DE cDNA encoding matrix metalloproteinase inhibitor, N-Tes.
XX KW Human; N-Tes; membrane-type matrix metalloproteinase inhibitor; cancer;
XX KW cancer infiltration; angiogenesis; Alzheimer's disease; arthritis;
XX KW glioma; cytostatic; antiarthritic; nootropic; neuroprotective;
XX KW antiangiogenic; gene; ss.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX CDS 108..1049
XX FT /*tag= a
XX FT /product= "N-Tes"
XX FT /note= "Matrix metalloproteinase inhibitor"
XX PN WO200257448-A1.
XX PD 25-JUL-2002.
XX PF 27-DEC-2001; 2001WO-JP011529.
XX PR 27-DEC-2000; 2000JP-00398817.
XX PA (DAII-) DAIICHI FINE CHEM CO LTD.
XX PI Sato H, Aoki T;
XX DR WPI; 2002-538473/57.
XX DR P-PSDB; AAU98506.
XX PT New human N-Tes polypeptide and its partial sequences inhibiting matrix
XX PT metalloproteinase for treatment and diagnosis of glioma and other
XX PT neurological and neoplastic disorders.
XX PS Claim 7; Page 114-115; 129pp; Japanese.
XX CC The invention describes polypeptides and their salts containing the
XX CC sequence of human N-Tes polypeptide or derived from it, and inhibiting
XX CC the activity of membrane-type matrix metalloproteinase. The polypeptides
XX CC are used for the treatment, prevention and diagnosis of cancer and cancer
XX CC infiltration, angiogenesis, Alzheimer's disease and arthritis, and
XX CC especially of glioma. This sequence encodes the novel N-Tes polypeptide
XX CC described in the invention
XX SQ Sequence 1510 BP; 507 A; 291 C; 335 G; 377 T; 0 U; 0 Other;

Query Match 1.3%; Score 48; DB 6; Length 1510;
Best Local Similarity 100.0%; Pred. No. 3.7e-07;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3715 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 1455 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1502

RESULT 84
AAA46483
ID AAA46483 standard; cDNA; 1544 BP.
XX AC AAA46483;
XX XX
XX DT 04-SEP-2000 (first entry)
XX DE cDNA encoding a filament-like protein 4.
XX KW MPPI binding factor 1; MAF1; transgenic plant; chimeric gene; NMPI;
XX KW nuclear matrix protein-1; FLIP; filament-like protein;
XX KW transcriptional activator; protein composition; crop growth;
XX KW crop protection; ss.
XX OS Lycopersicon esculentum.
XX KW

```

```

FH Key Location/Qualifiers
FT CDS 2..1021
FT /*tag= a
FT /product= "filament-like protein 4"
XX PN WO200028054-A2.
XX PD 18-MAY-2000.
XX PF 04-NOV-1999; 99WO-US025993.
XX PR 06-NOV-1998; 98US-00187999.
XX PA (DUPO) DU PONT DE NEMOURS & CO S I.
XX PI Gindullis F, Meier I;
XX DR WPI; 2000-376565/32.
XX DR P-PSDB; AAY93406.
XX PT Regulating gene expression in plant cells, useful e.g. for expressing new
XX PT traits, by introducing genes for transcriptional activators that interact
XX PT with nuclear matrix proteins.
XX PS Claim 24; Page 72; 82pp; English.
XX CC The present sequence encodes a filament-like protein 4 (FLIP4), which is
XX CC used in the course of the invention. The specification describes a method
XX CC whereby gene expression in a stably transformed transgenic plant cell is
XX CC regulated using a chimeric gene that encodes NMPI (nuclear matrix protein
XX CC -1) or FLIP. Two chimeric genes are combined into the genome, the first
XX CC gene comprising a promoter linked to a DNA binding domain, a coding
XX CC sequence or its complement, and a polyadenylation sequence. The second
XX CC gene comprises a promoter, a DNA binding domain, NMPI or FLIP
XX CC polynucleotides, and a polyadenylation sequence. Expression of the first
XX CC gene regulates expression of the second. NMPI, FLIP and related proteins
XX CC are parts of the nuclear skeleton and function as transcriptional
XX CC activators. The methods are used to regulate gene expression in plants,
XX CC e.g. to express novel traits in transgenic plants that may result in new
XX CC products useful in foods, pharmaceuticals and materials, or for
XX CC suppressing endogenous genes to alter the protein composition, and to
XX CC derive new phenotypes beneficial for crop growth and development. NMPI,
XX CC FLIP and related proteins are used to alter level of expression of MPPI,
XX CC binding factor 1 (MAPI) or MAF1-binding proteins in plants by co-
XX CC suppression or overexpression, or to screen for compounds that inhibit
XX CC activity of MPPI- or MAF1-binding proteins, which are potentially useful
XX CC as crop protection agents. They can also be used to study the plant
XX CC nucleus matrix
XX SQ Sequence 1544 BP; 492 A; 259 C; 367 G; 426 T; 0 U; 0 Other;

Query Match 1.3%; Score 48; DB 3; Length 1544;
Best Local Similarity 100.0%; Pred. No. 3.7e-07;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3715 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 1470 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1517

RESULT 85
ABK33555
ID ABK33555 standard; cDNA; 1837 BP.
XX AC ABK33555;
XX XX
XX DT 08-MAY-2002 (first entry)
XX DE cDNA encoding human PRO protein, Seq ID No 39.
XX KW Human; secreted protein; PRO; tumour; lung cancer; colon cancer;
XX KW breast cancer; prostate tumour; rectal tumour; liver tumour;
XX KW pericyte cell proliferation; chondrocyte cell proliferation;

```

tumour necrosis factor-alpha; gene; ss.

Homo sapiens.

WO200208288-A2.

31-JAN-2002.

29-JUN-2001; 2001WO-US021066.

20-JUL-2000; 2000US-0219556P.

25-JUL-2000; 2000US-0220585P.

25-JUL-2000; 2000US-0220605P.

25-JUL-2000; 2000US-0220607P.

25-JUL-2000; 2000US-0220624P.

25-JUL-2000; 2000US-0220638P.

25-JUL-2000; 2000US-0220654P.

25-JUL-2000; 2000US-0220666P.

26-JUL-2000; 2000US-0220893P.

28-JUL-2000; 2000WO-US020710.

01-AUG-2000; 2000US-0222425P.

22-AUG-2000; 2000US-0227133P.

23-AUG-2000; 2000WO-US023522.

24-AUG-2000; 2000WO-US023328.

10-NOV-2000; 2000WO-US030873.

28-NOV-2000; 2000US-0253646P.

01-DEC-2000; 2000WO-US032678.

20-DEC-2000; 2000US-00747259.

20-DEC-2000; 2000WO-US034956.

28-FEB-2001; 2001WO-US006520.

01-MAR-2001; 2001WO-US006666.

22-MAR-2001; 2001US-00816744.

10-MAY-2001; 2001US-00854208.

10-MAY-2001; 2001US-00854280.

25-MAY-2001; 2001WO-US017092.

(GETH) GENENTECH INC.

Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ,

Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;

WPI; 2002-172001/22.

P-PSDB; AAU83611.

One hundred and twenty two nucleic acids encoding PRO polypeptides,

useful for treating a PRO related disorder and for diagnosing tumors such

as lung cancer, colon cancer, breast tumor, prostate tumor, rectal tumor

or liver tumor.

Claim 2; Fig 39; 359pp; English.

The invention relates to one hundred and twenty two nucleic acids

encoding PRO polypeptides. The sequences of the 122 PRO polynucleotides

QY 3715 AATGAAA 3762
DB 1729 AATGAAA 1776

RESULT 86

ACA66860

ID ACA66860 standard; cDNA; 1837 BP.

XX ACA66860;

XX 23-JUN-2003 (first entry)

XX cDNA encoding human PRO polypeptide #20.

XX Human; PRO polypeptide; secreted and transmembrane protein;

XX anti-PRO antibody; diagnostic assay; gene expression; tumour; cytostatic;

XX Gene; ss.

XX Homo sapiens.

XX US2003036635-A1.

XX 20-FEB-2003.

XX 28-AUG-2002; 2002US-00230163.

XX 25-JUL-2000; 2000US-0220638P.

XX 01-JUN-2001; 2001WO-US017800.

XX 23-JUN-2001; 2001WO-US021066.

XX 03-APR-2002; 2002US-00119480.

XX (GETH) GENENTECH INC.

XX Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;

XX Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;

XX WPI; 2003-342045/32.

XX P-PSDB; ABU80758.

XX One hundred and twenty two nucleic acids encoding PRO polypeptides,

XX useful for the manufacture of a medicament for diagnosing or treating

XX tumor.

XX Claim 2; Fig 39; 314pp; English.

XX The present invention relates to the isolation of novel human PRO
XX polypeptides, and the polynucleotide sequences encoding them. The PRO
XX polypeptides are secreted and transmembrane proteins. The PRO
XX polypeptides and polynucleotides are useful for preparing a medicament
XX useful in the diagnosis and treatment of tumours. Anti-PRO antibodies are
XX useful in diagnostic assays for PRO, by detecting its expression in
XX specific cells, tissues or serum, and for affinity purification of PRO
XX from recombinant cell culture or natural sources. ACA66860-ACA66962
XX represent cDNA sequences encoding the human PRO polypeptides of the
XX invention. Note: The sequence data for this patent was obtained in
XX electronic format directly from the USPTO web site at
XX seqdata.uspto.gov/psipsDIDEntry.html

XX Sequence 1837 BP; 527 A; 454 C; 465 G; 391 T; 0 U; 0 Other;

XX Query Match 1.3%; Score 48; DB 7; Length 1837;

XX Best Local Similarity 100.0%; Pred. No. 3.6e-07;

XX Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAAA 3762

DB 1729 AATGAAA 1776

RESULT 87

ACD68612

Query Match 1.3%; Score 48; DB 6; Length 1837;

Best Local Similarity 100.0%; Pred. No. 3.6e-07;

Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

ID AC D68612 standard; cDNA; 1837 BP.
AC ACD68612;
XX
XX 17-SEP-2003 (first entry)
DT
XX
XX Novel human secreted and transmembrane protein PRO10111 cDNA.
DE
XX Human; secreted and transmembrane protein; PRO; cytostatic;
XX antiarthritic; osteopathic; gene therapy; TNF-Agonist-Alpha;
XX chondrocyte stimulator; pericyte stimulator; fibroblast modulator;
XX pharmaceutical; diagnostic; biosensor; bioreactor; tumour; lung tumour;
XX colon tumour; breast tumour; prostate tumour; rectal tumour;
XX liver tumour; bone disorder; cartilage disorder; sports injury;
XX arthritis; wound; gene; ss.
XX
XX Homo sapiens.
OS
XX US2003045687-A1.
XX
XX 06-MAR-2003.
PD
XX
XX 12-AUG-2002; 2002US-00218631.
XX
XX 01-JUN-2001; 2001WO-US017800.
XX
XX 29-JUN-2001; 2001WO-US021086.
XX
XX 09-APR-2002; 2002US-00119480.
XX
XX (GETH ) GENENTECH INC.
XX
XX Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;
XX Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;
XX
XX WPI; 2003-512315/48.
XX
XX P-PSDB; ABO33724.
XX
XX New genes, and its encoded secreted and transmembrane polypeptides,
XX useful for stimulating Tumor Necrosis Factor alpha, or chondrocyte or
XX pericyte proliferation, especially for treating lung tumors, arthritis or
XX wounds in a mammal.
XX
XX Claim 2; Fig 39; 314pp; English.
XX
XX The invention describes an isolated nucleic acid molecule comprising a
XX sequence with at least 80% identity to: (a) a nucleotide encoding any of
XX 122 PRO (secreted and transmembrane) polypeptides whose sequences are
XX fully defined in the specification; or (b) any of 122 nucleotide
XX sequences having e.g. 4834, 2504 or 1759 bp fully defined in the
XX specification; or the full length coding sequence of any these 122
XX nucleotide sequences. The PRO polypeptides or polynucleotides are useful
XX as pharmaceuticals, diagnostics, biosensors or bioreactors. These are
XX particularly useful for detecting tumours (e.g. lung tumour, colon
XX tumour, breast tumour, prostate tumour, rectal tumour, or liver tumour)
XX in a mammal, for stimulating the release of TNF-alpha from human blood,
XX for stimulating the proliferation or differentiation of chondrocyte
XX cells, for stimulating proliferation of pericyte cells, or for modulating
XX normal human dermal fibroblast proliferation. The PRO nucleic acid or
XX polypeptide is also useful for treating tumours or various bone and/or
XX cartilage disorders (e.g. sports injuries or arthritis), or wounds. The
XX PRO polypeptides are useful in drug screening, particularly as targets
XX for therapeutic intervention in these diseases, and in the diagnostic
XX determination of the presence of these diseases. The PRO polypeptides are
XX also useful as molecular weight markers, or for chromosome
XX screening libraries of human cDNA, genomic DNA or mRNA. The PRO genes may
XX also be used in gene therapy, particularly for replacing a defective
XX gene. This sequence encodes a novel human secreted and transmembrane PRO
XX polypeptide
XX
XX Sequence 1837 BP; 527 A; 454 C; 465 G; 391 T; 0 U; 0 Other;
XX
XX Query Match 1.3%; Score 48; DB 7; Length 1837;
XX Best Local Similarity 100.0%; Pred. No. 3.6e-07;
XX Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 3715 AATGAAAAA
DB 1729 AATGAAAAA

RESULT 89

ABT44245
ID ABT44245 standard; cDNA; 1837 BP.

AC ABT44245;
XX 06-NOV-2003 (first entry)
DT Human PRO10111 cDNA.

DE PRO; proliferation; pericyte cell; TNF-alpha; blood; chondrocyte; ss;
KW differentiation; dermal fibroblast; tumour; gene therapy; gene;
KW cytotstatic.

XX Homo sapiens.
OS
XX US2003050448-A1.

XX 13-MAR-2003.

XX 28-AUG-2002; 2002US-00230414.

XX 01-JUN-2001; 2001WO-US017800.
XX 29-JUN-2001; 2001WO-US021066.
XX 09-APR-2002; 2002US-00119480.

XX (GETH) GENENTECH INC.

XX Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;
PI Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;
PI WPI; 2003-521818/49.
XX P-PSDB; ABJ72247.

XX New nucleic acid encoding for a PRO protein, useful for the manufacture
PT of a medicament for diagnosing or treating tumors or for measuring or
PT detecting expression of an associated gene.

XX Claim 2; Fig 39; 315pp; English.

XX The invention relates to a novel isolated nucleic acid encoding a fully
CC defined PRO polypeptide. The molecules of the invention may be useful for
CC stimulating proliferation or gene expression in pericyte cells or the
CC release of TNF-alpha from human blood. Other possible uses include the
CC stimulation or inhibition of chondrocyte proliferation or
CC differentiation, the stimulation of human dermal fibroblast cell
CC proliferation and the detection of the presence of a tumour within a
CC mammal. Furthermore, the nucleic acid may be useful for the manufacture
CC of a medicament for diagnosing or treating a tumour within a mammal or
CC for measuring or detecting the expression of an associated gene, as well
CC as during gene therapy. The current sequence is that of the human PRO
CC cDNA of the invention

XX Sequence 1837 BP; 527 A; 454 C; 465 G; 391 T; 0 U; 0 Other;

Query Match 1.3%; Score 48; DB 8; Length 1837;
Best Local Similarity 100.0%; Fred. No. 3.6e-07;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAAAAA
DB 1729 AATGAAAAA

RESULT 90

ABT44528
ID ABT44528 standard; cDNA; 1837 BP.

XX

AC ABT44528;

XX 06-NOV-2003 (first entry)

XX Human PRO10111 cDNA.

XX PRO; proliferation; gene; pericyte cell; TNF alpha; chondrocyte; blood;
KW tumour necrosis factor; proliferation; differentiation; gene therapy;
KW dermal fibroblast; ss.

XX Homo sapiens.

XX US2003027988-A1.

XX 06-FEB-2003.

XX 26-AUG-2002; 2002US-00227884.

XX 01-JUN-2001; 2001WO-US017800.

XX 29-JUN-2001; 2001WO-US021066.

XX 09-APR-2002; 2002US-00119480.

XX (GETH) GENENTECH INC.

XX Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;
PI Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;
PI WPI; 2003-503301/47.
XX P-PSDB; ABJ72375.

XX New PRO protein encoding nucleic acid, useful for preparing PRO
PT polypeptides and anti-PRO antibodies for detecting the presence of a
PT tumor in a mammal.

XX Claim 2; Fig 39; 324pp; English.

XX The invention relates to a novel isolated PRO protein encoding nucleic
CC acid. The nucleic acid of the invention may be useful for preparing PRO
CC polypeptides and anti-PRO antibodies for detecting the presence of a
CC tumour in a mammal. Furthermore, the molecules of the invention may be
CC useful for stimulating proliferation or gene expression in pericyte
CC cells, the release of tumour necrosis factor (TNF)-alpha from human
CC blood, the proliferation or differentiation of chondrocyte cells and for
CC inhibiting the proliferation of normal human dermal fibroblast cells.
CC Finally, the molecules may be utilised during gene therapy. The current
CC sequence is that of the human PRO cDNA of the invention

XX Sequence 1837 BP; 527 A; 454 C; 465 G; 391 T; 0 U; 0 Other;

Query Match 1.3%; Score 48; DB 8; Length 1837;
Best Local Similarity 100.0%; Fred. No. 3.6e-07;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAAAAA
DB 1729 AATGAAAAA

RESULT 91

ACD82195
ID ACD82195 standard; cDNA; 1837 BP.

XX ACD82195;

XX 19-SEP-2003 (first entry)

XX Human secreted/transmembrane polypeptide PRO 10111 cDNA.

XX Human; ss; chondrocyte stimulation; TNF-alpha stimulation; gene therapy;
KW human dermal fibroblast stimulation; tumour; tissue typing; gene;
KW affinity purification.

XX Homo sapiens.

KW dermal fibroblast cell differentiation inhibitor; tumour; lung tumour;
KW colon tumour; breast tumour; prostate tumour; rectal tumour;
KW liver tumour; tissue typing; chromosome mapping; gene mapping;
KW gene therapy.
XX
XX Homo sapiens.
XX US2003073814-A1.
XX 17-APR-2003.
XX
XX 12-AUG-2002; 2002US-00218949.
XX
XX 01-JUN-2001; 2001WO-US017800.
XX 29-JUN-2001; 2001WO-US021066.
XX 09-APR-2002; 2002US-00119480.
XX
XX (GETH) GENENTECH INC.
XX
XX Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;
PI Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;
XX WPI; 2003-644806/61.
XX P-PSDB; ADB83530.
XX
XX New PRO polypeptides and nucleic acids encoding the polypeptides, useful
PT in gene therapy, chromosome identification, tissue typing, or as
PT hybridization probes in chromosome and gene mapping.
XX
XX Claim 2; Fig 39; 315pp; English.
XX
XX The invention describes an isolated PRO (secreted and transmembrane)
CC polypeptide (I). PRO982, PRO1160, PRO1187 or PRO1329 polypeptide are
CC useful for stimulating the proliferation of or gene expression in
CC pericyte cells. PRO357, PRO229, PRO1272 or PRO4405 polypeptide are useful
CC for stimulating the proliferation or differentiation of chondrocyte
CC cells. PRO231, PRO357, PRO725, PRO1155, PRO1306 or PRO1419 polypeptide
CC are useful for stimulating the release of tumour necrosis factor (TNF)-
CC alpha from human blood. PRO982, PRO357, PRO725, PRO1306, PRO1419, PRO214,
CC PRO247, PRO337, PRO526, PRO363, PRO331, PRO1083, PRO840, PRO1080,
CC PRO1478, PRO1134, PRO826, PRO1005, PRO809, PRO1071, PRO1411, PRO1309,
CC PRO1025, PRO1181, PRO1126, PRO1186, PRO1192, PRO1244, PRO1412,
CC PRO1286, PRO1330, PRO1347, PRO1273, PRO1279, PRO1340, PRO1338,
CC PRO1343, PRO1376, PRO1387, PRO1409, PRO1474, PRO1917, PRO1760, PRO1567,
CC PRO1887, PRO1928, PRO4341, PRO1801, PRO4333, PRO3543, PRO4344, PRO4332,
CC PRO9940, PRO6079, PRO9836 or PRO10096 polypeptide are useful for
CC stimulating the proliferation of normal human dermal fibroblasts cells.
CC
CC PRO181, PRO229, PRO788, PRO1194, PRO1272, PRO1488, PRO4302, PRO4408,
CC PRO5723, PRO5725, PRO7154, or PRO7425 polypeptide are useful for
CC inhibiting the proliferation of normal human dermal fibroblast cells. PRO
CC polypeptides such as PRO6004, PRO4981, PRO7174, PRO5778, PRO4332, etc.,
CC are useful for detecting the presence of tumour in a mammal which
CC involves comparing the level of expression of the above PRO polypeptides
CC in a test sample of cells taken from the mammal, and a control sample of
CC normal cells of the same cell type, where a higher level of expression of
CC the PRO polypeptides in the test sample as compared to the control sample
CC is indicative of the presence of tumour in the mammal. The tumour is lung
CC tumour, colon tumour, breast tumour, prostate tumour, rectal tumour or
CC liver tumour. (I) is useful as molecular weight markers, for tissue
CC typing, or as therapeutic agents. A polynucleotide (II) encoding (I) is
CC useful for chromosome and gene mapping or gene therapy. (II) is useful
CC for generating transgenic animals or knock-out animals which are useful
CC screening useful reagents. PRO357, PRO229, PRO1272 or PRO4405 polypeptide
CC is useful for treating bone and/or cartilage disorders (e.g., arthritis,
CC sport injuries). This sequence encodes a human secreted and transmembrane
CC PRO polypeptide.
XX
XX Sequence 1837 BP; 527 A; 454 C; 465 G; 391 T; 0 U; 0 Other;

Query Match 1.3%; Score 48; DB 8; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.6e-07;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAA 3762
|||||
DB 1729 AATGAA 1776
|||||
RESULT 94
ADB80635
ID ADB80635 standard; cDNA; 1837 BP.
XX
XX ADB80635;
XX
XX 04-DEC-2003 (first entry)
XX
XX Novel human secreted and transmembrane protein PRO10111 cDNA.
XX
XX Human; secreted and transmembrane protein; PRO; gene; ss; cytostatic;
KW vulnary; antiarthritic; pericyte cell proliferation;
KW pericyte cell differentiation; chondrocyte cell proliferation;
KW chondrocyte cell differentiation; tumour necrosis factor alpha release;
KW (TNF)-alpha release; dermal fibroblast cell proliferation; lung tumour;
KW dermal fibroblast cell differentiation inhibitor; tumour; lung tumour;
KW colon tumour; breast tumour; prostate tumour; rectal tumour;
KW liver tumour; tissue typing; chromosome mapping; gene mapping;
XX
XX Homo sapiens.
OS
XX US2003088068-A1.
XX
XX 08-MAY-2003.
XX
XX 13-AUG-2002; 2002US-00219481.
XX
XX 01-JUN-2001; 2001WO-US017800.
XX 29-JUN-2001; 2001WO-US021086.
XX 09-APR-2002; 2002US-00119480.
XX
XX (GETH) GENENTECH INC.
XX
XX Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;
PI Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;
XX WPI; 2003-657982/62.
XX P-PSDB; ADB80636.
XX
XX One hundred and twenty two nucleic acids encoding PRO polypeptides,
PT useful in gene therapy, chromosome identification, tissue typing, or as
PT hybridization probes in chromosome and gene mapping.
XX
XX Claim 2; Fig 39; 305pp; English.
XX
XX The invention describes an isolated PRO (secreted and transmembrane)
CC polypeptide (I). PRO982, PRO1160, PRO1187 or PRO1329 polypeptide are
CC useful for stimulating the proliferation of or gene expression in
CC pericyte cells. PRO357, PRO229, PRO1272 or PRO4405 polypeptide are useful
CC for stimulating the proliferation or differentiation of chondrocyte
CC cells. PRO331, PRO357, PRO725, PRO1155, PRO1306 or PRO1419 polypeptide
CC are useful for stimulating the release of tumour necrosis factor (TNF)-
CC alpha from human blood. PRO982, PRO357, PRO725, PRO1306, PRO1419, PRO214,
CC PRO247, PRO337, PRO526, PRO363, PRO331, PRO1083, PRO840, PRO1080,
CC PRO1478, PRO1134, PRO826, PRO1005, PRO809, PRO1071, PRO1411, PRO1309,
CC PRO1025, PRO1181, PRO1126, PRO1186, PRO1192, PRO1244, PRO1412,
CC PRO1286, PRO1330, PRO1347, PRO1273, PRO1279, PRO1340, PRO1338,
CC PRO1343, PRO1376, PRO1387, PRO1409, PRO1474, PRO1917, PRO1760, PRO1567,
CC PRO1887, PRO1928, PRO4341, PRO1801, PRO4333, PRO3543, PRO4344, PRO4332,
CC PRO9940, PRO6079, PRO9836 or PRO10096 polypeptide are useful for
CC stimulating the proliferation of normal human dermal fibroblasts cells.
CC
CC PRO181, PRO229, PRO788, PRO1194, PRO1272, PRO1488, PRO4302, PRO4408,
CC PRO5723, PRO5725, PRO7154, or PRO7425 polypeptide are useful for
CC inhibiting the proliferation of normal human dermal fibroblast cells. PRO
CC polypeptides such as PRO6004, PRO4981, PRO7174, PRO5778, PRO4332, etc.,
CC are useful for detecting the presence of tumour in a mammal which
CC involves comparing the level of expression of the above PRO polypeptides
CC in a test sample of cells taken from the mammal, and a control sample of
CC normal cells of the same cell type, where a higher level of expression of
CC the PRO polypeptides in the test sample as compared to the control sample
CC is indicative of the presence of tumour in the mammal. The tumour is lung
CC tumour, colon tumour, breast tumour, prostate tumour, rectal tumour or
CC liver tumour. (I) is useful as molecular weight markers, for tissue
CC typing, or as therapeutic agents. A polynucleotide (II) encoding (I) is
CC useful for chromosome and gene mapping or gene therapy. (II) is useful
CC for generating transgenic animals or knock-out animals which are useful
CC screening useful reagents. PRO357, PRO229, PRO1272 or PRO4405 polypeptide
CC is useful for treating bone and/or cartilage disorders (e.g., arthritis,
CC sport injuries). This sequence encodes a human secreted and transmembrane
CC PRO polypeptide.
XX
XX Sequence 1837 BP; 527 A; 454 C; 465 G; 391 T; 0 U; 0 Other;

CC in a test sample of cells taken from the mammal, and a control sample of
 CC normal cells of the same cell type, where a higher level of expression of
 CC the PRO polypeptides in the test sample as compared to the control sample
 CC is indicative of the presence of tumour in the mammal. The tumour is lung
 CC tumour, colon tumour, breast tumour, prostate tumour, rectal tumour or
 CC liver tumour. (I) is useful as molecular weight markers, for tissue
 CC typing, or as therapeutic agents. A polynucleotide (II) encoding (I) is
 CC useful for chromosome and gene mapping or gene therapy. (II) is useful
 CC for generating transgenic animals or knock-out animals which are useful
 CC screening useful reagents. PRO357, PRO229, PRO1272 or PRO4405 polypeptide
 CC is useful for treating bone and/or cartilage disorders (e.g., arthritis,
 CC sport injuries). This sequence encodes a human secreted and transmembrane
 CC PRO polypeptide.

XX SQ Sequence 1837 BP; 527 A; 454 C; 465 G; 391 T; 0 U; 0 Other;

Query Match 1.3%; Score 48; DB 8; Length 1837;
 Best Local Similarity 100.0%; Pred. No. 3.6e-07;
 Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAAA 3762

DB 1729 AATGAAA 1776

RESULT 95

ID ADB73176 standard; cDNA; 1837 BP.

AC ADB73176;

DT 04-DEC-2003 (first entry)

XX DE Novel human secreted and transmembrane protein PRO10111 cDNA.

XX human, secreted and transmembrane protein; PRO; gene; ss; cytostatic;
 KW vulnary; antiarthritic; pericyte cell proliferation;
 KW pericyte cell differentiation; chondrocyte cell proliferation;
 KW chondrocyte cell differentiation; tumour necrosis factor alpha release;
 KW (TNF)-alpha release; dermal fibroblast cell proliferation;
 KW dermal fibroblast cell differentiation inhibitor; tumour; lung tumour;
 KW colon tumour; breast tumour; prostate tumour; rectal tumour;
 KW liver tumour; tissue typing; chromosome mapping; gene mapping;
 KW gene therapy.

XX OS Homo sapiens.

XX PN US2003096968-A1.

XX PD 22-MAY-2003.

XX PF 29-AUG-2002; 2002US-00232223.

XX PR 01-JUN-2001; 2001WO-US017800.

XX PR 29-JUN-2001; 2001WO-US021066.

XX PR 09-APR-2002; 2002US-00119480.

XX PA (GETH) GENENTECH INC.

XX PI Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;

XX PI Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;

XX DR WPI; 2003-765525/72.

XX DR P-PSDB; ADB73177.

XX XX New isolated PRO polypeptides useful as molecular weight markers in

PT protein electrophoresis, useful for tissue typing, and for treating

PT arthritis and tumors.

XX XX Claim 2; Fig 39; 308pp; English.

PS The invention describes an isolated PRO (secreted and transmembrane)

CC polypeptide (I). PRO982, PRO1160, PRO1187 or PRO1329 polypeptide are

CC useful for stimulating the proliferation of or gene expression in
 CC pericyte cells. PRO357, PRO229, PRO1272 or PRO4405 polypeptide are useful
 CC for stimulating the proliferation or differentiation of chondrocyte
 CC cells. PRO357, PRO725, PRO1155, PRO1306 or PRO1419 polypeptide
 CC are useful for stimulating the release of tumour necrosis factor (TNF)-
 CC alpha from human blood. PRO982, PRO357, PRO1306, PRO1419, PRO214,
 CC PRO247, PRO337, PRO526, PRO363, PRO531, PRO1083, PRO840, PRO1080,
 CC PRO1478, PRO1134, PRO826, PRO1005, PRO809, PRO1071, PRO1309,
 CC PRO1025, PRO1181, PRO1126, PRO1186, PRO1192, PRO1244, PRO1412,
 CC PRO1286, PRO1330, PRO1347, PRO1305, PRO1279, PRO1340, PRO1338,
 CC PRO1343, PRO1376, PRO1387, PRO1409, PRO1474, PRO1917, PRO1760, PRO1567,
 CC PRO1887, PRO1928, PRO341, PRO1801, PRO4333, PRO3543, PRO4322,
 CC PRO940, PRO6079, PRO9836 or PRO10096 polypeptide are useful for
 CC stimulating the proliferation of normal human dermal fibroblasts cells.
 CC PRO181, PRO229, PRO788, PRO1194, PRO1272, PRO1488, PRO4302, PRO4408,
 CC PRO5723, PRO5725, PRO7154, or PRO7425 polypeptide are useful for
 CC inhibiting the proliferation of normal human dermal fibroblast cells. PRO
 CC polypeptides such as PRO6004, PRO4981, PRO7174, PRO5778, PRO4332, etc.,
 CC are useful for detecting the presence of tumour in a mammal which
 CC involves comparing the level of expression of the above PRO polypeptides
 CC in a test sample of cells taken from the mammal, and a control sample of
 CC normal cells of the same cell type, where a higher level of expression of
 CC the PRO polypeptides in the test sample as compared to the control sample
 CC is indicative of the presence of tumour in the mammal. The tumour is lung
 CC tumour, colon tumour, breast tumour, prostate tumour, rectal tumour or
 CC liver tumour. (I) is useful as molecular weight markers, for tissue
 CC typing, or as therapeutic agents. A polynucleotide (II) encoding (I) is
 CC useful for chromosome and gene mapping or gene therapy. (II) is useful
 CC for generating transgenic animals or knock-out animals which are useful
 CC screening useful reagents. PRO357, PRO229, PRO1272 or PRO4405 polypeptide
 CC is useful for treating bone and/or cartilage disorders (e.g., arthritis,
 CC sport injuries). This sequence encodes a human secreted and transmembrane
 CC PRO polypeptide.

XX SQ Sequence 1837 BP; 527 A; 454 C; 465 G; 391 T; 0 U; 0 Other;

Query Match 1.3%; Score 48; DB 8; Length 1837;

Best Local Similarity 100.0%; Pred. No. 3.6e-07;

Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAAA 3762

DB 1729 AATGAAA 1776

RESULT 96

ID ADB78258

XX ADB78258 standard; cDNA; 1837 BP.

XX AC ADB78258;

XX DT 04-DEC-2003 (first entry)

XX DE Novel human secreted and transmembrane protein PRO10111 cDNA.

XX human, secreted and transmembrane protein; PRO; gene; ss; cytostatic;
 KW vulnary; antiarthritic; pericyte cell proliferation;
 KW pericyte cell differentiation; chondrocyte cell proliferation;
 KW chondrocyte cell differentiation; tumour necrosis factor alpha release;
 KW (TNF)-alpha release; dermal fibroblast cell proliferation;
 KW dermal fibroblast cell differentiation inhibitor; tumour; lung tumour;
 KW colon tumour; breast tumour; prostate tumour; rectal tumour;
 KW liver tumour; tissue typing; chromosome mapping; gene mapping;
 KW gene therapy.

XX OS Homo sapiens.

XX PN US2003092889-A1.

XX PD 15-MAY-2003.

XX PF 13-AUG-2002; 2002US-00219478.

ID ADB78012 standard; cDNA; 1837 BP.
AC ADB78012;
XX
DT 04-DEC-2003 (first entry)
XX
DE Novel human secreted and transmembrane protein PRO10111 cDNA.
XX
KW Human; secreted and transmembrane protein; PRO; gene; ss; cytostatic;
KW vulnary; antiarthritic; pericyte cell proliferation;
KW pericyte cell differentiation; chondrocyte cell proliferation;
KW chondrocyte cell differentiation; tumour necrosis factor alpha release;
KW (TNF)-alpha release; dermal fibroblast cell proliferation;
KW dermal fibroblast cell differentiation inhibitor; tumour; lung tumour;
KW colon tumour; breast tumour; prostate tumour; rectal tumour;
KW liver tumour; tissue typing; chromosome mapping; gene mapping;
KW gene therapy.
XX
OS Homo sapiens.
XX
XX US2003092886-A1.
XX
XX 15-MAY-2003.
XX
XX 09-AUG-2002; 2002US-00216165.
XX
XX 25-JUL-2000; 2000US-0220607P.
XX
XX 01-JUN-2001; 2001WO-US017800.
XX
XX 29-JUN-2001; 2001WO-US021066.
XX
XX 09-APR-2002; 2002US-00119480.
XX
XX (GETH) GENENTECH INC.
XX
XX Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;
XX Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;
XX WPI; 2003-765494/72.
XX
XX P-PSDB; ADB78013.
XX
XX Novel isolated PRO polypeptide useful for tissue typing, gene therapy, as
XX molecular weight markers in protein electrophoresis, for treating
XX arthritis, tumor.
XX
XX Claim 2; Fig 39; 308pp; English.
XX
XX The invention describes an isolated PRO (secreted and transmembrane)
XX polypeptide (I). PRO982, PRO1160, PRO1187 or PRO1329 polypeptide are
XX useful for stimulating the proliferation of or gene expression in
XX pericyte cells. PRO357, PRO229, PRO1272 or PRO4405 polypeptide are useful
XX for stimulating the proliferation or differentiation of chondrocyte
XX cells. PRO231, PRO357, PRO725, PRO155, PRO1306 or PRO1419 polypeptide
XX are useful for stimulating the release of tumour necrosis factor (TNF)-
XX alpha from human blood. PRO982, PRO357, PRO725, PRO1306, PRO1419, PRO214,
XX PRO247, PRO337, PRO526, PRO363, PRO531, PRO1083, PRO840, PRO1080,
XX PRO1478, PRO1134, PRO826, PRO1005, PRO809, PRO1071, PRO1411, PRO1309,
XX PRO1025, PRO1124, PRO1126, PRO1186, PRO1192, PRO1274, PRO1412,
XX PRO1285, PRO1330, PRO1347, PRO1305, PRO1273, PRO1279, PRO1340, PRO1338,
XX PRO1345, PRO1376, PRO1387, PRO1409, PRO1474, PRO1917, PRO1760, PRO1567,
XX PRO1887, PRO1928, PRO4341, PRO1801, PRO4333, PRO3543, PRO3444, PRO4332,
XX PRO9940, PRO6079, PRO9836 or PRO10096 polypeptide are useful for
XX stimulating the proliferation of normal human dermal fibroblasts cells.
XX PRO181, PRO229, PRO788, PRO1194, PRO1272, PRO1488, PRO4302, PRO4408,
XX PRO5723, PRO5725, PRO7154, or PRO7425 polypeptide are useful for
XX inhibiting the proliferation of normal human dermal fibroblast cells. PRO
XX polypeptides such as PRO6004, PRO4981, PRO1714, PRO5778, PRO4332, etc.,
XX are useful for detecting the presence of tumour in a mammal which
XX involves comparing the level of expression of the above PRO polypeptides
XX in a test sample of cells taken from the mammal, and a control sample of
XX normal cells of the same cell type, where a higher level of expression of
XX the PRO polypeptides in the test sample as compared to the control sample
XX is indicative of the presence of tumour in the mammal. The tumour is lung
XX tumour, colon tumour, breast tumour, prostate tumour, rectal tumour or
XX liver tumour. (I) is useful as molecular weight markers, for tissue

CC typing, or as therapeutic agents. A polynucleotide (II) encoding (I) is
CC useful for chromosome and gene mapping or gene therapy. (II) is useful
CC for generating transgenic animals or knock-out animals which are useful
CC screening useful reagents. PRO357, PRO229, PRO1272 or PRO4405 polypeptide
CC is useful for treating bone and/or cartilage disorders (e.g., arthritis,
CC sport injuries). This sequence encodes a human secreted and transmembrane
CC PRO polypeptide.
XX
SQ Sequence 1837 BP; 527 A; 454 C; 465 G; 391 T; 0 U; 0 Other;
Query Match 1.3%; Score 48; DB 9; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.6e-07;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3715 AATGAAA 3762
DB 1729 AATGAAA 1776

RESULT 99
ADB78078
ID ADB87078 standard; cDNA; 1837 BP.
XX
XX ADB87078;
XX
XX 04-DEC-2003 (first entry)
XX
XX Human PRO polynucleotide #20.
XX
XX Human; PRO; gene; ss; secreted polypeptide; transmembrane polypeptide;
KW tumour; cancer; lung; colon; breast; prostate; rectum; liver;
KW tumour necrosis factor-alpha; TNF-alpha; blood; chondrocyte cell;
KW pericyte cell; dermal fibroblast; bone disorder; cartilage disorder;
KW arthritis; sports injury; cytostatic; antiarthritic.
XX
XX Homo sapiens.
XX
XX US2003088067-A1.
XX
XX 08-MAY-2003.
XX
XX 13-AUG-2002; 2002US-00219479.
XX
XX 01-JUN-2001; 2001WO-US017800.
XX
XX 29-JUN-2001; 2001WO-US021066.
XX
XX 09-APR-2002; 2002US-00119480.
XX
XX (GETH) GENENTECH INC.
XX
XX Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;
XX Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;
XX WPI; 2003-657981/62.
XX
XX P-PSDB; ADB87079.
XX
XX One hundred and twenty two nucleic acids encoding PRO polypeptides,
XX useful in gene therapy, chromosome identification, tissue typing, or as
XX hybridization probes in chromosome and gene mapping.
XX
XX Claim 2; Fig 39; 314pp; English.
XX
XX The invention relates to human PRO polypeptides (secreted and
XX transmembrane polypeptides) and the PRO polynucleotides encoding them.
XX The PRO polypeptides and polynucleotides are useful as pharmaceuticals,
XX diagnostics, biosensors or bioreactors. They are particularly useful for
XX detecting tumours (e.g. lung tumour, colon tumour, breast tumour,
XX prostate tumour, rectal tumour or liver tumour) in a mammal, for
XX stimulating the release of tumour necrosis factor (TNF)-alpha from human
XX blood, for stimulating the proliferation or differentiation of
XX chondrocyte cells, for stimulating the proliferation of or gene
XX expression in pericyte cells or for stimulating the proliferation of
XX normal human dermal fibroblasts. The PRO nucleic acids are useful as
XX hybridisation probes, in chromosome and gene mapping, in generating

antisense RNA and DNA, in preparing PRO polypeptides by recombinant technology, in generating transgenic animals or knock-out animals which may be used in the development and screening of therapeutically useful reagents, in gene therapy, in chromosome identification, as chromosome markers and in generating probes. The PRO polypeptides, or anti-PRO antibodies, are useful for preparing a medicament for treating a condition which is responsive to the PRO polypeptides or anti-PRO antibodies, such as pericyte-associated tumours and bone and/or cartilage disorders (e.g. arthritis, sports injuries), involving inducing the re-differentiation of chondrocytes. The PRO polypeptides are useful as molecular markers for protein electrophoresis, and in tissue typing. This sequence represents a human PRO polynucleotide of the invention.

XX SQ Sequence 1837 BP; 527 A; 454 C; 465 G; 391 T; 0 U; 0 Other;

Query Match 1.3%; Score 48; DB 9; Length 1837;

Best Local Similarity 100.0%; Pred. No. 3.6e-07;

Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAAAAA 3762

DB 1729 AATGAAAAA 1776

RESULT 100

ID ADB84660 standard; cDNA; 1837 BP.

AC ADB84660;

DT 04-DEC-2003 (first entry)

XX Human PRO polynucleotide #20.

XX Human; PRO; gene; ss; secreted polypeptide; transmembrane polypeptide; tumour; cancer; lung; colon; breast; prostate; rectum; liver; tumour necrosis factor-alpha; TNF-alpha; blood; chondrocyte cell; pericyte cell; dermal fibroblast; bone disorder; cartilage disorder; arthritis; sports injury; cytostatic; antiarthritic.

XX Homo sapiens.

XX US2003092890-A1.

XX 15-MAY-2003.

XX 14-AUG-2002; 2002US-00219536.

XX 28-JUL-1999; 99US-0146222P.

XX 24-FEB-2000; 2000WO-US005004.

XX 02-MAR-2000; 2000WO-US005841.

XX 01-JUN-2001; 2001WO-US017800.

XX 29-JUN-2001; 2001WO-US021066.

XX 09-APR-2002; 2002US-00119480.

XX (GETH) GENENTECH INC.

XX Baker KP, Deenoyers L, Gerritsen ME, Goddard A, Godowski PJ;

XX Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;

XX WPI; 2003-777259/73.

XX P-PSDB; ADB84661.

XX New isolated PRO polypeptides, useful for tissue typing, gene therapy, as molecular weight markers in protein electrophoresis, and for treating arthritis and tumors.

XX Claim 2; Fig 39; 308pp; English.

XX The invention relates to human PRO polypeptides (secreted and transmembrane polypeptides) and the PRO polynucleotides encoding them. The PRO polypeptides and polynucleotides are useful as pharmaceuticals, diagnostics, biosensors or bioreactors. They are particularly useful for

CC detecting tumours (e.g. lung tumour, colon tumour, breast tumour, CC prostate tumour, rectal tumour or liver tumour) in a mammal, for CC stimulating the release of tumour necrosis factor (TNF)-alpha from human CC blood, for stimulating the proliferation or differentiation of CC chondrocyte cells, for stimulating the proliferation of or gene CC expression in pericyte cells or for stimulating the proliferation of CC normal human dermal fibroblasts. The PRO nucleic acids are useful as CC hybridisation probes, in chromosome and gene mapping, in generating CC antisense RNA and DNA, in preparing PRO polypeptides by recombinant CC technology, in generating transgenic animals or knock-out animals which CC may be used in the development and screening of therapeutically useful CC reagents, in gene therapy, in chromosome identification, as chromosome CC markers and in generating probes. The PRO polypeptides, or anti-PRO CC antibodies, are useful for preparing a medicament for treating a CC condition which is responsive to the PRO polypeptides or anti-PRO CC antibodies, such as pericyte-associated tumours and bone and/or cartilage CC disorders (e.g. arthritis, sports injuries), involving inducing the re- CC differentiation of chondrocytes. The PRO polypeptides are useful as CC molecular markers for protein electrophoresis, and in tissue typing. This CC sequence represents a human PRO polynucleotide of the invention.

XX SQ Sequence 1837 BP; 527 A; 454 C; 465 G; 391 T; 0 U; 0 Other;

Query Match 1.3%; Score 48; DB 9; Length 1837;

Best Local Similarity 100.0%; Pred. No. 3.6e-07;

Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAAAAA 3762

DB 1729 AATGAAAAA 1776

Search completed: April 23, 2004, 04:26:38

Job time : 1354 secs.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 23, 2004, 04:04:06 ; Search time 8944 Seconds
(without alignments)
12560.546 Million cell updates/sec

Title: -US-10-005-907-1

Perfect score: 3762

Sequence: 1 gagaaaccgactgtgta.....aaaaaaaaaaaaaaaaaaaaa 3762

Scoring table: OLIGO_NUC
Gapop_60.0 , Gapext 60.0

Searched: 27513289 seqs, 14931090276 residues

Word size: 0

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 150 summaries

Database: EST.*

- 1: em_estba.*
- 2: em_esthum.*
- 3: em_estin.*
- 4: em_estmu.*
- 5: em_estov.*
- 6: em_estpl.*
- 7: em_estro.*
- 8: em_estci.*
- 9: gb_estl.*
- 10: gb_est2.*
- 11: gb_est3.*
- 12: gb_est4.*
- 13: gb_est5.*
- 14: gb_estfun.*
- 15: em_eston.*
- 16: em_eston.*
- 17: em_gss_hum.*
- 18: em_gss_inv.*
- 19: em_gss_pln.*
- 20: em_gss_vit.*
- 21: em_gss_fun.*
- 22: em_gss_mus.*
- 23: em_gss_mus.*
- 24: em_gss_pro.*
- 25: em_gss_rod.*
- 26: em_gss_phg.*
- 27: em_gss_vrl.*
- 28: gb_gss1.*
- 29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	665	17.7	741	13 BU15309	BU15309 AGENCOURT
2	624	16.6	785	13 BU171872	BU171872 AGENCOURT
3	620	16.5	782	13 BO441475	BO441475 AGENCOURT
4	602	16.0	728	14 CD639433	CD639433 AGENCOURT

5	595	15.8	749	13 BU852381	BU852381
6	556	14.8	704	28 AQ390121	AQ390121 CITBI-E1-
7	530	14.1	549	10 BF434786	BF434786 7072F03.X
8	506	13.5	527	10 BF242113	BF242113 60180401
9	471	12.5	757	28 BZ600841	BZ600841 WHAAUJ99TR
10	462	12.3	786	13 BU155429	BU155429 AGENCOURT
11	441	11.7	504	10 AW665849	AW665849 h194F08.X
12	438	11.6	456	9 AI766257	AI766257 wh8904.X
13	387	10.3	549	12 BG461807	BG461807 RST44735
14	316	8.4	416	9 AA447527	AA447527 zw81b09.8
15	233	6.2	385	28 AQ006627	AQ006627 CIT-HSP-2
16	194	5.2	809	10 BF240893	BF240893 601875444
17	163	4.3	848	10 BF243193	BF243193 601877334
18	80	2.1	643	29 AG182348	AG182348 Pan tlogi
19	53	1.4	712	12 EG779038	EG779038 602667021
20	52	1.4	344	14 CA301835	CA301835 tao08910.
21	52	1.4	733	9 AV706915	AV706915 AV706915
22	51	1.4	190	14 CA780165	CA780165 MPL384_4
23	51	1.4	322	14 CA779205	CA779205 MPL384_8
24	51	1.4	502	13 BX512688	BX512688 BX512688
25	51	1.4	649	13 EQ571861	EQ571861 UI-M-PCO-
26	51	1.4	931	13 BX370381	BX370381 BX370381
27	51	1.4	1201	13 BX387156	CF328728 NACL--03-
28	50	1.3	80	14 CF929023	CF929023 laf76f02.
29	50	1.3	142	14 CF929023	CF929023 laf76f02.
30	50	1.3	146	14 CF755431	CF755431 lae60c03.
31	50	1.3	148	14 CD645057	CD645057 lab18e09.
32	50	1.3	148	14 CF259903	CF259903 lab41h07.
33	50	1.3	149	14 CF545723	CF545723 lac99f06.
34	50	1.3	152	14 CF970419	CF970419 lag03a12.
35	50	1.3	156	14 CF382304	CF382304 lac61h06.
36	50	1.3	158	14 CF929504	CF929504 laf79a10.
37	50	1.3	171	10 AW315672	AW315672 13217 MAR
38	50	1.3	174	14 CF354150	CF354150 lab94b10.
39	50	1.3	175	12 BM177915	BM177915 saj67a02.
40	50	1.3	183	14 CF979915	CF979915 lac74b06.
41	50	1.3	183	14 CF970684	CF970684 lag79a01.
42	50	1.3	208	14 CF970587	CF970587 lag77f05.
43	50	1.3	215	14 CF425041	CF425041 lad14d09.
44	50	1.3	230	14 CF352096	CF352096 lab55c12.
45	50	1.3	233	9 AL045558	AL045558 DKF2p434L
46	50	1.3	249	14 CF928508	CF928508 lae52b08.
47	50	1.3	250	14 CF805201	CF805201 lac64c06.
48	50	1.3	255	9 AW022602	AW022602 df41b12.Y
49	50	1.3	259	10 BE332081	BE332081 ut07g01.Y
50	50	1.3	261	10 AW733655	AW733655 sk83d01.Y
51	50	1.3	261	14 CD645268	CD645268 lab20b12.
52	50	1.3	265	10 AW827290	AW827290 xnl1e09.Y
53	50	1.3	269	14 CF352983	CF352983 lab60b08.
54	50	1.3	293	14 CF380398	CF380398 lac46f08.
55	50	1.3	293	14 CF380717	CF380717 lac50f08.
56	50	1.3	294	14 CF929133	CF929133 laf77h05.
57	50	1.3	301	14 CF384267	CF384267 lac04h02.
58	50	1.3	305	9 AL045620	AL045620 DKF2p434O
59	50	1.3	314	14 CF754869	CF754869 lae07d02.
60	50	1.3	321	14 CD422638	CD422638 lae83g11.
61	50	1.3	329	9 AU034841	AU034841 AU034841
62	50	1.3	336	13 BU767819	BU767819 SURBDC02
63	50	1.3	346	14 CF382593	CF382593 lac88g07.
64	50	1.3	350	13 BO450975	BO450975 PFESTOABO
65	50	1.3	354	10 BE048163	BE048163 tz46g05.Y
66	50	1.3	366	13 BM306217	BM306217 BM306217
67	50	1.3	387	12 BG958555	BG958555 359509 MA
68	50	1.3	395	14 CF979559	CF979559 rg78c04.Y
69	50	1.3	407	9 AA493109	AA493109 rg96602.8
70	50	1.3	427	9 AI133559	AI133559 HA2177 Hu
71	50	1.3	440	14 CF621874	CF621874 lae68d10.
72	50	1.3	441	14 CF979574	CF979574 rg83e02.Y
73	50	1.3	443	12 BM958905	BM958905 PLATE 11
74	50	1.3	457	13 CS1260	CS1260 C91260 Dict
75	50	1.3	464	9 AV707816	AV707816 AV707816
76	50	1.3	464	14 CF546023	CF546023 lad9a06.
77	50	1.3	477	14 CB524236	CB524236 UI-M-GK0-

REFERENCE 1 (bases 1 to 782)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: CLONTECH
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L10W2334 row: h column: 12
High quality sequence stop: 561.
Location/Qualifiers
1..782
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6101435"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_82"
/note="Organ: testis; Vector: pDNR-LIB (Clontech); Site: 1: SfiI (ggccattatggcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGGCCATTATGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCGGCGGCGGCACATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.35 kb (range 0.9-4.0 kb). 14/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

ORIGIN
Query Match 16.5%; Score 620; DB 13; Length 782;
Best Local Similarity 100.0%; Pred. No. 1.5e-108; Mismatches 0; Indels 0; Gaps 0;
Matches 620; Conservative 0;
3029 GGAACTTTGAATCCAGAGCATCAATCTCTCTGGTGGTTCCACCAATAGCCACAGCAGA 3088
3030 GGAACTTTGAATCCAGAGCATCAATCTCTCTGGTGGTTCCACCAATAGCCACAGCAGA 62
3089 TGTCTTAATCTTCCGAGATCTAGTTTTCAGCAAGCAGGATTAGAAATGTAACAT 3148
63 TGTCTTAATCTTCCGAGATCTAGTTTTCAGCAAGCAGGATTAGAAATGTAACAT 122
3149 CTTATGTGTTATGAAGAACATAGAAATCATCTGCTATAGTGTCTTTTAACTGTAAA 3208
123 CTTATGTGTTATGAAGAACATAGAAATCATCTGCTATAGTGTCTTTTAACTGTAAA 182
3209 TTTTGTGAAGCTATCTTTTATGCATATAAATATTGGAACATTTTACATTTTATATTT 3268
183 TTTTGTGAAGCTATCTTTTATGCATATAAATATTGGAACATTTTACATTTTATATTT 242
3269 TTAATCAGTTTCTCAAGTGTGATTATATATACAGAAATATGACCTGTAAAGGTAGA 3328
243 TTAATCAGTTTCTCAAGTGTGATTATATATACAGAAATATGACCTGTAAAGGTAGA 302
3329 GTTATAGAAATTTGTCAAAATGATTTCCCAATGATGACCTCTTATGAAGACAGA 3388
303 GTTATAGAAATTTGTCAAAATGATTTCCCAATGATGATGACCTCTTATGAAGACAGA 362
3389 ACAGTGTACATCTCCAGAAAGTTCACAGTGTCTCTTTTCCCTGAGTTTCACAGTCT 3448
363 ACAGTGTACATCTCCAGAAAGTTCACAGTGTCTCTTTTCCCTGAGTTTCACAGTCT 422
3449 GCGAACCAATGATCTGCTTCGATATAATTAATTAATTAATTAATTAATTAATTAATTA 3508
423 GCGAACCAATGATCTGCTTCGATATAATTAATTAATTAATTAATTAATTAATTAATTA 482
3509 CTTTCCATATTTATTTTGTGTGTGAAGCTCTTTTGTAGTCTTTTATATATATTTTGAT 3568

Db 483 CTTTCCATATTTATTTTGTGTGTGAAGCTCTTTTATAGTCATATAATATTTTGAGAT 542
QY 3569 TCATCTATGTTTAAATGTTCTATCATCAGTAGTGTACATCTTACTTGTCTCAGCATATCACCA 3628
Db 543 TCATCTATGTTTAAATGTTCTATCATCAGTAGTGTACATCTTACTTGTCTCAGCATATCACCA 602
QY 3629 TATAGTATATATATATTTG 3648
Db 603 TATAGTATATATATATTTG 622
RESULT 4
CD639433 728 bp mRNA linear EST 17-JUN-2003
AGENCOURT_14532881 NIH_MGC_191 Homo sapiens cDNA clone
IMAGE:30414375 5', mRNA sequence.
CD639433
ACCESSION CD639433.1 GI:31806240
VERSION EST.
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 728)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-remail.nih.gov
Tissue Procurement: Narayan Bhat
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: NDCM210 row: n column: 16
High quality sequence stop: 604.
Location/Qualifiers
1..728
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30414375"
/issue_type="Pooled"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_191"
/note="Vector: pDNR-LIB; Site 1: SfiI (ggccattatggcc); Site 2: SfiI (ggccgctcggcc); Library is oligo-dT primed and directionally cloned. PBMC - Peripheral Blood Mononuclear Cells. RNA was pooled from 3/6hour stimulation with PMA and Ionomycin. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGGCCATTATGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCGGCGGCGGCACATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.69 kb (range 0.70-5.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC Library."

FEATURES
source
Query Match 16.0%; Score 602; DB 14; Length 728;
Best Local Similarity 100.0%; Pred. No. 4e-105;
Matches 602; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1835 ATCTCTACCAAGGTGGATGTGTAGTTTCTGCTTTTAAATTTCAAGCAAACTGGAATA 1894
Db 3 ATCTCTACCAAGGTGGATGTGTAGTTTCTGCTTTTAAATTTCAAGCAAACTGGAATA 62

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 704)
AUTHORS Zhao,S., Adams,M.D., Nierman,W., Malek,J., Shizuya,H., Simon,M. and Venter,J.C.
TITLE Use of BAC End Sequences from Caltech Libraries for Sequence-Ready Map Building
JOURNAL Unpublished (1997)
COMMENT Other_GSSs: CITBI-EL-2548B21.TR
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbe@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tldb/hungen/bac_end_search/bac_end_search.html.
Seq primer: M3-21
Class: BAC ends.

FEATURES
source
1..704
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="2548B21"
/sex="male"
/cell_type="sperm"
/clone_lib="CITBI-EL"
/note="Vector: pBeloBAC11; Site_1: EcoRI; Site_2: EcoRI; Caltech Human BAC Library D"

ORIGIN
Query Match 14.8%; Score 556; DB 28; Length 704;
Best Local Similarity 99.7%; Pred. No. 2e-96;
Matches 656; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 803 GAAATTCGGCCATGATAGGAGGAGGTGACACACACTTGTATACCCCTTCCTTTG 862
DB 6 GAAATTCGGCCATGATAGGAGGAGGTGACACACACTTGTATACCCCTTCCTTTG 65
QY 863 GAGTTTATGCACAGTGACAGGATGAGTCTATAGACTGATGAATAGACTGTTGGC 922
DB 66 GAGTTTATGCACAGTGACAGGATGAGTCTATAGACTGATGAATAGACTGTTGGC 125
QY 923 AATAGAGTCCCAATTCACACTGCTGCTGTAGATCACACACTGTCTGAGGGATTCC 982
DB 126 AATAGAGTCCCAATTCACACTGCTGCTGTAGATCACACACTGTCTGAGGGATTCC 185
QY 983 ATCTATGAGACTTTGTCTACATAACAGAGACTTGGTTTCCAAACCCCTTATTTAGC 1042
DB 186 ATCTATGAGACTTTGTCTACATAACAGAGACTTGGTTTCCAAACCCCTTATTTAGC 245
QY 1043 TAAAGCATTTTCTACTGACTTCTTAAGCTTTAGACAAAGCTTAACCTTTCAACCA 1102
DB 246 TAAAGCATTTTCTACTGACTTCTTAAGCTTTAGACAAAGCTTAACCTTTCAACCA 305
QY 1103 ATTGCCAATCAGACAACTTTGAATCTACCTATGACCTGTGAAGCTCTCTCTGCTTCAAG 1162
DB 306 ATTGCCAATCAGACAACTTTGAATCTACCTATGACCTGTGAAGCTCTCTCTGCTTCAAG 365
QY 1163 ATCTTGCTCTTAAAGCTGAAACCGATGCACTTCCATTTAATGATTTATGCTTTGCT 1222
DB 366 ATCTTGCTCTTAAAGCTGAAACCGATGCACTTCCATTTAATGATTTATGCTTTGCT 425
QY 1223 TGTAACTCTGCTCCCTAAATGTATAAAGTAAACCGTCACTGACCTGACCACTCAGGCAC 1282
DB 426 TGTAACTCTGCTCCCTAAATGTATAAAGTAAACCGTCACTGACCACTCAGGCAC 485
QY 1283 ACTTCTCAGGACCTCTGAGGTGATCCAGGCCATGGTAAAGTCATGTTGGCTCAGAA 1342

DB 486 ACTTCTCAGGACCTCTCAGAGTGATATCCAGGGCCATGGTAAGTCATGTTGGCTCAGAA 545
QY 1343 TCAACCTCTTTAAATATTTTACAGAAATTTGGTTTGGTTTACCAATAAGTCTCCCAAAAT 1402
DB 546 TCAACCTCTTTAAATATTTTACAGAAATTTGGTTTGGTTTACCAATAAGTCTTCACAAAT 605
QY 1403 ATATGTCCAGAAATCTTCAATTCAGGCTGCTCACCAGAAATTTCAATGCCAATCT 1460
DB 606 ATATGTCCAGAAATCTTCAATTCAGGCTGCTCACCAGAAATTTCAATGCCAATCT 663

RESULT 7
BF434786/c
LOCUS BF434786
DEFINITION mRNA sequence.
ACCESSION BF434786
VERSION BF434786
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 549)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-romail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL, send email to: info@image.llnl.gov
Seq primer: -40UP from Gibco
High quality sequence stop: 476.

FEATURES
source
1..549
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3641981"
/lab_host="DH10B"
/clone_lib="NCI CGAP Kid11"
/note="Organ: Kidney; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker, site 1: Not I; Site 2: Eco RI; Plasmid DNA from the normalized library NCI_CGAP Kid3 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneID# 132376-132391, 1456007-1456775, and 1500552-1502855). Subtraction by Bento Soares and M. Fatima Bonaldo."

ORIGIN
Query Match 14.1%; Score 530; DB 10; Length 549;
Best Local Similarity 100.0%; Pred. No. 2.1e-91;
Matches 530; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3191 TCGTTTTTAACTGTAAATTTTGTGAAGCTTATCTTTTATGCATATAATATTTCACAT 3250
DB 549 TCGTTTTTAACTGTAAATTTTGTGAAGCTTATCTTTTATGCATATAATATTTCACAT 490
QY 3251 TTATACCTGTTTATATTTTAAATCAGTTTTTACTCAAGTGTGATTATATACAGAAATGT 3310
DB 489 TTATACCTGTTTATATTTTAAATCAGTTTTTACTCAAGTGTGATTATATACAGAAATGT 430
QY 3311 AACCACTGTAAAGGTAGAGTTATAGAAATTTTGTCAAATGTATTCCACCATGTAGTCACC 3370

```

Db 429 AACCACTGTAGGTTATAGAAATTTTGCAATGTATTCACCCATGTAGTCACC 370
QY 3371 TCCTTATGAGAGACAGAACACGTACATCTCCAGAAAGTTCACAGTGTCTCTTTTC 3430
Db 369 TCCTTATGAGAGACAGAACACGTACATCTCCAGAAAGTTCACAGTGTCTCTTTTC 310
QY 3431 CTGAGTTTCACAGTCTCTGGCAACCAATGATCTCTCGTATATATATACTGTCTCAGA 3490
Db 309 CTGAGTTTCACAGTCTCTGGCAACCAATGATCTCTCGTATATATATACTGTCTCAGA 250
QY 3491 TATTTGTAGCAATGTACCTTTCCATATTTATTTTGTGTGTAGGCTTCTTTTAGTCA 3550
Db 249 TATTTGTAGCAATGTACCTTTCCATATTTATTTTGTGTGTAGGCTTCTTTTAGTCA 190
QY 3551 TTATAATATTTTTCAGATTCATCTATGTTTAATGTTCTATCAGTAGTGTGACATCTTACT 3610
Db 189 TTATAATATTTTTCAGATTCATCTATGTTTAATGTTCTATCAGTAGTGTGACATCTTACT 130
QY 3611 TGTCTCAGCATATCACCATATAGATATATATAATTTTGTTAATCTAATCACTGATGATA 3670
Db 129 TGTCTCAGCATATCACCATATAGATATATATAATTTTGTTAATCTAATCACTGATGATA 70
QY 3671 TGTAGGATATTTAGTTTTCACATATATGAATAAGTGGCTATATAATGAA 3720
Db 69 TGTAGGATATTTAGTTTTCACATATATGAATAAGTGGCTATATAATGAA 20

RESULT 8
BF242113 527 bp mRNA linear EST 14-NOV-2000
LOCUS 601880401F1 NIH_MGC_55 Homo sapiens cDNA clone IMAGE:4109064 5',
DEFINITION mRNA sequence.
ACCESSION BF242113
VERSION BF242113.1 GI:11156040
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 527)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: CLONETECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCW993 row: a column: 01
High quality sequence stop: 518.
Location/Qualifiers
1. 527
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4109064"
/tissue_type="from acute myelogenous leukemia"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH MGC 55"
/note="Organ: bone marrow; Vector: pDNR-LIB (Clontech);
Site 1: SfiI (ggcgctcgcc); Site 2: SfiI
(ggcatatggcc); Double-stranded cDNA was prepared from
cell line RNA. 5' and 3' adaptors were used in cloning as
follows: 5' adaptor sequence: 5'-CACGGCCATTATGCCC-3' and
3' adaptor sequence:
5'-ATCTAGAGCGCGCGCGGCACATG-dT(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size
1.65 kb (range 0.9-4.0 kb). 14/15 colonies contained

```

inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

```

ORIGIN
Query Match 13.5%; Score 506; DB 10; Length 527;
Best Local Similarity 100.0%; Pred. No. 7.4e-87;
Matches 506; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGAAACGAGTCTGTGAAAGATGGAAATATATCTCTCGGAAACTCAGTTGCCGTG 60
Db 21 GAGAAACGAGTCTGTGAAAGATGGAAATATATCTCTCGGAAACTCAGTTGCCGTG 80
QY 61 GGAGAGAAATCAAAAGAACCCCAAGAAAGGAAACCCAGATGAGGAAAGAAACCGCAGGAA 120
Db 81 GGAGAGAAATCAAAAGAACCCCAAGAAAGGAAACCCAGATGAGGAAAGAAACCGCAGGAA 140
QY 121 ATGACTACATTTTGAAGAAACTTCAAGATCAAGATAGAAAGCCAAAGAGTTTCATCC 180
Db 141 ATGACTACATTTTGAAGAAACTTCAAGATCAAGATAGAAAGCCAAAGAGTTTCATCC 200
QY 181 ACTTCTTAATCAGAAACCCAGAAATGGCAGTGGTCTTGAAGAGTGTGCTACATCTCAT 240
Db 201 ACTTCTTAATCAGAAACCCAGAAATGGCAGTGGTCTTGAAGAGTGTGCTACATCTCAT 260
QY 241 AATCAGATCCCCCATCAGAGATCTCCCTGAGCTCAATGATGATGGCTATGAGACATT 300
Db 261 AATCAGATCCCCCATCAGAGATCTCCCTGAGCTCAATGATGATGGCTATGAGACATT 320
QY 301 GACTCCCTCACAAGGAAAGTGAAGACAGATTTAGAGAAAGTTCAGACACAGAAATATGCCCTT 360
Db 321 GACTCCCTCACAAGGAAAGTGAAGACAGATTTAGAGAAAGTTCAGACACAGAAATATGCCCTT 380
QY 361 CTTAGAGCTTCTGTAGTAGGCTTGTCTCCGACCCATGACATGATTCAGATTG 420
Db 381 CTTAGAGCTTCTGTAGTAGGCTTGTCTCCGACCCATGACATGATTCAGATTG 440
QY 421 TTTCACACATAAATCTCAAGCTGTCTTATCACCTTCACCAATGAAGACAAATGCAGAA 480
Db 441 TTTCACACATAAATCTCAAGCTGTCTTATCACCTTCACCAATGAAGACAAATGCAGAA 500
QY 481 TAGCAGACTCTGGCGAAGTTGTTCAC 506
Db 501 TAGCAGACTCTGGCGAAGTTGTTCAC 526

```

```

RESULT 9
BF2600841 757 bp DNA linear GSS 08-JUN-2003
LOCUS WHAAJ89TR Human MCF7 breast cancer cell line library (MCF7_1) Homo
DEFINITION sapiens genomic clone MCF7_1-3010, genomic survey sequence.
ACCESSION BF2600841
VERSION BF2600841.1 GI:31509303
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 757)
AUTHORS Volik S., Zhao S., Chin K., Brebner, J.H., Herndon, D.R., Tao, Q.,
Kowbel, D., Huang G., Lapuk, A., Kuo, W.-L., Magrane, G., de Jong, P.,
Gray, J.W., and Collins, C.
TITLE End-sequence profiling: Sequence-based analysis of aberrant genomes
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 100 (13), 7696-7701 (2003)
MEDLINE 22709111
PUBMED 12788976
COMMENT Contact: Volik SV
Colin Collins' lab
UCSF Comprehensive Cancer Center
UCSF Box 0808, San Francisco, CA 94143-0808, USA
Tel: 415 502 7066
Fax: 415 502 5665
Email: svolik@cc.ucsf.edu

```

This clone is available from Amplicon Express
http://www.genomex.com
Class: BAC ends.

FEATURES

source
Location/Qualifiers
1..757
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="MCF7_1-3010"
/sex="female"
/clone_lib="Human MCF7 breast cancer cell line library (MCF7_1)"
/note="Vector: pECBAC1; Site 1: HindIII; This library was constructed from MCF7 breast cancer cell line by Amplicon Express (http://www.genomex.com) using their standard procedure."

ORIGIN

Query Match 12.5%; Score 471; DB 28; Length 757;
Best Local Similarity 100.0%; Pred. No. 2.1e-80; Mismatches 0; Indels 0; Gaps 0;
Matches 471; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3222 ATCTTTTATGCATATAATTTGAACATTTTACATTTGTTTATATTTTATCATGTTTGA 3281
Db 1 ATCTTTTATGCATATAATTTGAACATTTTACATTTGTTTATATTTTATCATGTTTGA 60
QY 3282 CTCAGTCTGATTATATACAGAAATGTAACCACTGTAGGCTAGAGTTATAGAAATTT 3341
Db 61 CTCAGTCTGATTATATACAGAAATGTAACCACTGTAGGCTAGAGTTATAGAAATTT 120
QY 3342 TGTCAATGTATTCACCCATGTAGTCACCTCTTATGAAGAGACAGACAGTACATCCT 3401
Db 121 TGTCAATGTATTCACCCATGTAGTCACCTCTTATGAAGAGACAGACAGTACATCCT 180
QY 3402 CCCAGAAAGTCCACAGTCTCTTTCCCTGAGTTTCCACAGTCTCTGGGAAACCAATGAT 3461
Db 181 CCCAGAAAGTCCACAGTCTCTTTCCCTGAGTTTCCACAGTCTCTGGGAAACCAATGAT 240
QY 3462 CTGCTTCGTATATTAATTAACCTGTTCTAGATATTTGTAGCAATGTACCTTTCCATATTTA 3521
Db 241 CTGCTTCGTATATTAATTAACCTGTTCTAGATATTTGTAGCAATGTACCTTTCCATATTTA 300
QY 3522 TTTTGTGTGTGAAGCTCTTTTATGTCATTAATATTTTGTAGATTCATCTATGTTTGA 3581
Db 301 TTTTGTGTGTGAAGCTCTTTTATGTCATTAATATTTTGTAGATTCATCTATGTTTGA 360
QY 3582 ATGTTCTATCATGATGTGTACATCTTCTGTCCTCAGCATATCACCATATAGATATACTA 3641
Db 361 ATGTTCTATCATGATGTGTACATCTTCTGTCCTCAGCATATCACCATATAGATATACTA 420
QY 3642 TAAATTTGTTAATCTAATCACTGATGATATGATGATATTTAAGTTTGTGA 3692
Db 421 TAAATTTGTTAATCTAATCACTGATGATATGATGATATTTAAGTTTGTGA 471

RESULT 10

BU155429
LOCUS BU155429
DEFINITION AGENCOURT_7979368 NIH_MGC_55 Homo sapiens cDNA clone IMAGE:6108338
5', mRNA sequence.
ACCESSION BU155429
VERSION BU155429.1 GI:22668961
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 786)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
CONTACT: Robert Strausberg, Ph.D.
Email: cgapbs-romail.nih.gov

Tissue Procurement: ATCC
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at:
http://image.llnl.gov
Plate: LLCM2352
High quality sequence stop: 399.
Location/Qualifiers
1..786
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6108338"
/tissue_type="from acute myelogenous leukemia"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH MGC 55"
/note="Organ: bone marrow; Vector: pDNR-LIB (Clontech); Site 1: SfiI (ggcgctcgcc); Site 2: SfiI (ggcattatggcc); Double-stranded cDNA was prepared from cell line RNA. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGCCCATTAATGCCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCGGCGCGCCGACATG-DT(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.65 kb (range 0.9-4.0 kb). 14/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

ORIGIN

Query Match 12.3%; Score 462; DB 13; Length 786;
Best Local Similarity 99.8%; Pred. No. 1e-78; Mismatches 1; Indels 0; Gaps 0;
Matches 512; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2609 GTTTCAGATTTTGGAGCATTTTCAGATCTTCAGATTTAGGATTTTCAACTGTACTGACCT 2668
Db 3 GTTTCAGATTTTGGAGCATTTTCAGATCTTCAGATTTAGGATTTTCAACTGTACTGACCT 62
QY 2669 TTTAGTCATTGACAGCATTAATCAATAGTGGATCCAGATTAATCTATTTGCTGTATAC 2728
Db 63 TTTAGTCATTGACAGCATTAATCAATAGTGGATCCAGATTAATCTATTTGCTGTATAC 122
QY 2729 ACATTTTGCTCTCTATTCAACGAAATCTTATSCCCTCTTGTGTGATTTTAAATGTGCGG 2788
Db 123 ACATTTTGCTCTCTATTCAACGAAATCTTATSCCCTCTTGTGTGATTTTAAATGTGCGG 182
QY 2789 AAGGGAACAATAGAAATTTTGCATTTCTAGAAAGTCATTTCTGCAAAATATGTCAGTC 2848
Db 183 AAGGGAACAATAGAAATTTTGCATTTCTAGAAAGTCATTTCTGCAAAATATGTCAGTC 242
QY 2849 CTGTAGATATTAGCCAAATTTTAGGAAAATGACAAAATTTTACCTTTTCTGCTGCTTTG 2908
Db 243 CTGTAGATATTAGCCAAATTTTAGGAAAATGACAAAATTTTACCTTTTCTGCTGCTTTG 302
QY 2909 TAGCTGTTTATGATATAAATACCTTTTGTATAAATAATTAATTTTAAATTTAGTACACA 2968
Db 303 TAGCTGTTTATGATATAAATACCTTTTGTATAAATAATTAATTTTAAATTTAGTACACA 362
QY 2969 ATCTGGAATTTATCAGAGAAGGGGCAAGCAATAGGTTTATAAACAAGTATTGTTGGTAGAA 3028
Db 363 ATCTGGAATTTATCAGAGAAGGGGCAAGCAATAGGTTTATAAACAAGTATTGTTGGTAGAA 422
QY 3029 GGAACCTTGAATCCAGAGCATCAATGCTTCTGCTGCTGTTCCATTAAGCCACAGCAGA 3088
Db 423 GGAACCTTGAATCCAGAGCATCAATGCTTCTGCTGCTGTTCCATTAAGCCACAGCAGA 482
QY 3089 TGCTTAACTCTTCCGAGATCTAGTTTTTTCAGC 3121
Db 483 TGCTTAACTCTTCCGAGATCTAGTTTTTTCAGC 515

```

RESULT 11
AW65849/c      AW65849      504 bp      mRNA      linear      EST 06-APR-2000
LOCUS          h194f08.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
DEFINITION     IMAGE:2979975 3', mRNA sequence.
ACCESSION      AW65849
VERSION        AW65849.1 GI:7458398
KEYWORDS       EST.
SOURCE         Homo sapiens (human)
ORGANISM       Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 504)
AUTHORS        NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE          National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
              Tumor Gene Index
JOURNAL        Unpublished (1997)
COMMENT        Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-remail.nih.gov
              This clone is available royalty-free through LLNL; contact the
              IMAGE Consortium (info@image.llnl.gov) for further information.
              Seq primer: -40UP from Gibco
              High quality sequence stop: 451.
              Location/Qualifiers
                1..504
                  /organism="Homo sapiens"
                  /mol_type="mRNA"
                  /db_xref="taxon:9606"
                  /clone="IMAGE:2979975"
                  /lab_host="DH10B"
                  /clone_lib="Soares_NFL_T_GBC_S1"
                  /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
                  a modified polylinker; Site 1: Not 1; Site 2: Eco RI;
                  Equal amounts of plasmid DNA from three normalized
                  libraries (fetal lung NBH19w, testis NHT, and B-cell
                  NCI CGAP-SCB1) were mixed and ss circles were made in
                  vitro. Following HAP purification, this DNA was used as
                  tracer in a subtractive hybridization reaction. The driver
                  was PCR-amplified cDNAs from pools of 5,000 clones made
                  from the same 3 libraries. The pools consisted of
                  I.M.A.G.E. clones 297480-302087, 682632-687239,
                  726408-728711, and 729096-731399. Subtraction by Bento
                  Soares and M. Fatima Bonaldo."
FEATURES       source
               11.7%; Score 441; DB 10; Length 504;
               Best Local Similarity 99.8%; Pred. No. 1.5e-74;
               Matches 491; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
ORIGIN
Query Match      11.7%; Score 441; DB 10; Length 504;
Best Local Similarity 99.8%; Pred. No. 1.5e-74;
Matches 491; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 3219 CTTATCTTTTATGCATATAAATATTGACATTTTACATGCTTTTATATTTTATATCGATT 3278
Db 504 CTTATCTTTTATGCATATAAATATTGACATTTTACATGCTTTTATATTTTATATCGATT 445
Qy 3279 TTACTCAAGTGTGATTATATACAGAAATGTAACCACTGTAAGGGGTAGAGTTATAGAA 3338
Db 444 TTACTCAAGTGTGATTATATACAGAAATGTAACCACTGTAAGGGGTAGAGTTATAGAA 395
Qy 3339 TTTTGTCAAATGTATTACCCATGTATGTCACCTCTTATGAGAGACAGAACACGTACAT 3398
Db 384 TTTTGTCAAATGTATTACCCATGTATGTCACCTCTTATGAGAGACAGAACACGTACAT 325
Qy 3399 CTTCCAGAAAGTTCACAGTGTCTTTTCCCTGAGTTTCACAGTCTCTGGCAACCAAT 3458
Db 324 CTTCCAGAAAGTTCACAGTGTCTTTTCCCTGAGTTTCACAGTCTCTGGCAACCAAT 265
Qy 3459 GATCTGCTCGTATTAATTAATTAACGTTCAGATATTTGTAGCAATATGACCTTTCCATAT 3518
Db 264 GATCTGCTCGTATTAATTAATTAACGTTCAGATATTTGTAGCAATATGACCTTTCCATAT 205
Qy 3519 TTATTTCTGCTGTAAGCGTCTCTTTTAGTCATTATATATTTTGTAGATTCATCTATGT 3578
Db 204 TTATTTCTGCTGTAAGCGTCTCTTTTAGTCATTATATATTTTGTAGATTCATCTATGT 145

```

```

Qy 3579 TTAATGTTCTATCAGTAGTCTTACATCTTCTCTCAGCATATCACCATATAGATATA 3638
Db 144 TTAATGTTCTATCAGTAGTCTTACATCTTCTCTCAGCATATCACCATATAGATATA 85
Qy 3639 CTATAATTTGTTAATCTAATCACTGATGATATAGGATATTTAAGTTTTTGACATTAT 3698
Db 84 CTATAATTTGTTAATCTAATCACTGATGATATAGGATATTTAAGTTTTTGACATTAT 25
Qy 3699 GAATAAAGTGGC 3710
Db 24 GAATAAAGTGGC 13
RESULT 12
AW66257/c      AW66257      456 bp      mRNA      linear      EST 20-DEC-1999
LOCUS          wh8604.x1 NCI_CGAP_Kid1l Homo sapiens cDNA clone IMAGE:2385942 3',
DEFINITION     mRNA sequence.
ACCESSION      AW66257
VERSION        AW66257.1 GI:5232766
KEYWORDS       EST.
SOURCE         Homo sapiens (human)
ORGANISM       Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 456)
AUTHORS        NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE          National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
              Tumor Gene Index
JOURNAL        Unpublished (1997)
COMMENT        Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-remail.nih.gov
              Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
              Emmert-Buck, M.D., Ph.D.
              cDNA Library Preparation: M. Bento Soares, Ph.D.
              cDNA Library Arrayed by: Greg Lennon, Ph.D.
              DNA Sequencing by: Washington University Genome Sequencing Center
              Clone distribution: NCI-CGAP clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              www-bio.llnl.gov/bbrp/image/image.html
              Insert Length: 1086 Std Error: 0.00
              Seq primer: -40UP from Gibco
              High quality sequence stop: 437.
              Location/Qualifiers
                1..456
                  /organism="Homo sapiens"
                  /mol_type="mRNA"
                  /db_xref="taxon:9606"
                  /clone="IMAGE:2385942"
                  /lab_host="DH10B"
                  /clone_lib="NCI_CGAP_Kid1l"
                  /note="Organ: Kidney; Vector: pT7T3D-Pac (Pharmacia) with
                  a modified polylinker; Site 1: Not 1; Site 2: Eco RI;
                  Plasmid DNA from the normalized library NCI_CGAP Kid3 was
                  prepared, and ss circles were made in vitro. Following HAP
                  purification, this DNA was used as tracer in a subtractive
                  hybridization reaction. The driver was PCR-amplified cDNAs
                  from a pool of 5,000 clones made from the same library
                  (cloneIDB 1322376-1323911, 1456007-1456775, and
                  1500552-1502855). Subtraction by Bento Soares and M.
                  Fatima Bonaldo."
FEATURES       source
               11.6%; Score 438; DB 9; Length 456;
               Best Local Similarity 100.0%; Pred. No. 6e-74;
               Matches 438; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
ORIGIN
Query Match      11.6%; Score 438; DB 9; Length 456;
Best Local Similarity 100.0%; Pred. No. 6e-74;
Matches 438; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3283 TCAAGTGTGATTATATACAGAAATGTAAACCACTGTAGGGTAGATTATAGAAATTTT 3342
Db 456 TCAAGTGTGATTATATACAGAAATGTAAACCACTGTAGGGTAGATTATAGAAATTTT 397
Qy 3343 GTCAATGTATTACCCATGTAGTCACCTCTTATGAGAGACAGAACACGTACATCTCTC 3402

```


Db 396 GTCRAATGATATCCACCATGTAGTCACTCCTTATGAAGACAGACAGATACATCCTC 337
 QY 3403 CCAGAAAGTTCACAGTCTCTCTTTCCCTGAGTTTACAGTCTCTGCGACCAATGATC 3462
 Db 336 CCAGAAAGTTCACAGTCTCTCTTTCCCTGAGTTTACAGTCTCTGCGACCAATGATC 277
 QY 3463 TGCCTTCGTAATATTAATTAATCTGTTCTAGATATTTGTAGCAATGTACCTTTCCCATATTTAT 3522
 Db 276 TGCCTTCGTAATATTAATTAATCTGTTCTAGATATTTGTAGCAATGTACCTTTCCCATATTTAT 217
 QY 3523 TTTGTGTGTGAAGCTCTCTTTTGTAGTCAATTAATATTTTGTAGTCAATTAATTTAA 3582
 Db 216 TTTGTGTGTGAAGCTCTCTTTTGTAGTCAATTAATATTTTGTAGTCAATTAATTTAA 157
 QY 3583 TGTTCATACAGTAGTTGACATCTTACTTGTCTCAGGATATCACCATATAGATATCTAT 3642
 Db 156 TGTTCATACAGTAGTTGACATCTTACTTGTCTCAGGATATCACCATATAGATATCTAT 97
 QY 3643 AATTTGTTAACTAATCACTGATGATATGTAGGATATTTAAAGTTTGTGACATTAATGAAT 3702
 Db 96 AATTTGTTAACTAATCACTGATGATATGTAGGATATTTAAAGTTTGTGACATTAATGAAT 37
 QY 3703 AAAGTGGCTATAAATGAA 3720
 Db 36 AAAGTGGCTATAAATGAA 19

RESULT 13
 BG461807/c
 LOCUS RST44735 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
 DEFINITION BG461807
 ACCESSION BG461807.1 GI:13750313
 VERSION EST.
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
 Cain,S., Leventhal,C., Thornton,M., Ramachandran,K.,
 Whittington,J., Lerner,L., Costanzo,D., McElligott,K., Booser,S.,
 Mays,R., Smith,E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K.,
 Offenbacher,J., Danzig,J. and Ducar M.
 Creation of genome-wide protein expression libraries using random
 activation of gene expression
 Nat. Biotechnol. 19 (5), 440-445 (2001)

JOURNAL
 MEDLINE
 PUBMED
 COMMENT
 Contact: Scott J. Cain
 Athersys, Inc.
 3201 Carnegie Ave, Cleveland, OH 44115, USA
 Tel: 216 431 9900
 Fax: 216 361 9596
 Email: scain@atersys.com.
 Location/Qualifiers
 1..549
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /cell_line="HT1080"
 /clone_lib="Athersys RAGE Library"
 /note="See 'Creation of Genome-wide Protein Expression
 Libraries using Random Activation of Gene Expression',
 Nature Biotechnology, in press. Note that even though the
 cell type indicated is HT1080, since a random activation
 method was used, these sequence tags are not necessarily
 expressed in HT1080 under normal circumstances."

FEATURES
 source
 1..549
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /cell_line="HT1080"
 /clone_lib="Athersys RAGE Library"
 /note="See 'Creation of Genome-wide Protein Expression
 Libraries using Random Activation of Gene Expression',
 Nature Biotechnology, in press. Note that even though the
 cell type indicated is HT1080, since a random activation
 method was used, these sequence tags are not necessarily
 expressed in HT1080 under normal circumstances."

ORIGIN
 Query Match 10.3%; Score 387; DB 12; Length 549;
 Best Local Similarity 100.0%; Pred. No. 2.1e-64;
 Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3316 CTGTAAGGTAGAGTATTAAGAAATTTTGTCAAATGTAATTCACCCATGTAGTCACTCCTT 3375
 Db 387 CTGTAAGGTAGAGTATTAAGAAATTTTGTCAAATGTAATTCACCCATGTAGTCACTCCTT 328
 QY 3376 ATGAAGAGACAGAACACGCTACATCTCCAGAAAGTTCACAGTCTCTCTTTCCCTGAG 3435
 Db 327 ATGAAGAGACAGAACACGCTACATCTCCAGAAAGTTCACAGTCTCTCTTTCCCTGAG 268
 QY 3436 TTTTACCAGTCTCTGCGAACCAATGATCTGCTCGTATTAATTAATTAATTAATTAATTT 3495
 Db 267 TTTTACCAGTCTCTGCGAACCAATGATCTGCTCGTATTAATTAATTAATTAATTT 208
 QY 3496 GTAGCAATGTACCTTTCCATATTTTGTGTGTGAAGGCTTCTTTTAGTCAATTATA 3555
 Db 207 GTAGCAATGTACCTTTCCATATTTTGTGTGTGAAGGCTTCTTTTAGTCAATTATA 148
 QY 3556 ATATTTTTCAGATTCATCTATCTTTTAAATGTTCTATCAGTCTGTGACATCTTACTGTCT 3615
 Db 147 ATATTTTTCAGATTCATCTATCTTTTAAATGTTCTATCAGTCTGTGACATCTTACTGTCT 88
 QY 3616 CAGCATATCACCATATAGATATCTATTAATTTGTTAACTTAATCTAATCAGTATGTAG 3675
 Db 87 CAGCATATCACCATATAGATATCTATTAATTTGTTAACTTAATCTAATCAGTATGTAG 28
 QY 3676 GATATTTAAAGTTTTCAGATTAATGAAT 3702
 Db 27 GATATTTAAAGTTTTCAGATTAATGAAT 1

RESULT 14
 AA447527/c
 LOCUS zw81b09.sl Soares testis NHT Homo sapiens cDNA clone IMAGE:782585
 DEFINITION 3', mRNA sequence.
 ACCESSION AA447527
 VERSION AA447527.1 GI:21611197
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
 Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J.,
 Moore,B., Schellenberg,K., Steptoe,M., Tan,P., Theising,B.,
 White,Y., Wyllie,T., Waterston,R. and Wilson,R.
 WashU-Merck EST Project 1997
 Unpublished (1997)
 Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@wustl.wustl.edu
 This clone is available royalty-free through LNL ; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Seq primer: -41m13 fwd. Et from Amersham
 High quality sequence stop: 102.

TITLE
 JOURNAL
 COMMENT
 Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@wustl.wustl.edu
 This clone is available royalty-free through LNL ; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Seq primer: -41m13 fwd. Et from Amersham
 High quality sequence stop: 102.

FEATURES
 source
 1..416
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:782585"
 /sex="male"
 /lab_host="DH10B"
 /clone_lib="Soares testis NHT"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
 was prepared from mRNA obtained from Clontech
 Laboratories, Inc., and primed with a Not I - oligo(dT)
 primer [5]

5'-ATTCTAGAGCGCGCGCGCATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.65 kb (range 0.9-4.0 kb). 14/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

ORIGIN

Query Match 5.2%; Score 194; DB 10; Length 809;
Best Local Similarity 99.5%; Pred. No. 4e-28;
Matches 364; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 718 AATGAAATTAATTTCTAATCCCTGACTAATGATGACCCCTCTTAGGCCAAAG 777
DB 137 AATGAAATTAATTTCTAATCCCTGACTAATGATGACCCCTCTTAGGCCAAAG 196

QY 778 AGACCTCAGATGAACCTGAAGACTGAATTTGGCCATGATAGGAGGTGAGACAC 837
DB 197 AGACCTCAGATGAACCTGAAGACTGAATTTGGCCATGATAGGAGGTGAGACAC 256

QY 838 ACCTTGTTATACCCCTTCCCTTTGGAGTTTATGACAAAGTGACAGGATGATGAAG 897
DB 257 ACCTTGTTATACCCCTTCCCTTTGGAGTTTATGACAAAGTGACAGGATGATGAAG 315

QY 898 ACTGATGAATAGACTGATTTGGCAATAGAGTCCCAATCCACCTGACTCTGGTGTA 957
DB 316 ACTGATGAATAGACTGATTTGGCAATAGAGTCCCAATCCACCTGACTCTGGTGTA 375

QY 958 GATCACACACTGTCTGAGGAGTTCCATCTATGAGACTTTGTCTACATAACAGACCTTG 1017
DB 376 GATCACACACTGTCTGAGGAGTTCCATCTATGAGACTTTGTCTACATAACAGACCTTG 435

QY 1018 GTTCCACACACCCCTTATTTTAGCTAAGACTTTTCTACTGACTCTTAGTCTTT 1077
DB 436 GTTCCACACACCCCTTATTTTAGCTAAGACTTTTCTACTGACTCTTAGTCTTT 1077

QY 1078 AGACAA 1083
DB 496 AGACAA 501

RESULT 17

BF243193
LOCUS 601877334F1 NIH_MGC_55 Homo sapiens cDNA clone IMAGE:4105665 5',
DEFINITION mRNA sequence. 848 bp mRNA linear EST 14-NOV-2000

ACCESSION BF243193

VERSION BF243193.1 GI:11157122

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 848)

AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

TITLE NIH-MGC http://mgi.nci.nih.gov/

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: CLONETECH Laboratories, Inc.

DNA Sequencing by: Incyte Genomics, Inc.

Clone Distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLCM984 row: c column: 10

High quality sequence stop: 546.

Location/Qualifiers

1. .848

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

FEATURES

source

Vector: pBACe3.6

R.site 1: EcorI

R.site 2: EcorI

Location/Qualifiers

1. .643

/organism="Pan troglodytes"

/clone="IMAGE:4105665"
/tissue type="from acute myelogenous leukemia"
/lab host="DH10B (T1 phage-resistant)"
/clone lib="NIH MGC 55"
/note="Organ: bone marrow; Vector: pDNR-LIB (Clontech);
Site 1: SfiI (ggccattatggcc); Site 2: SfiI (ggccattatggcc); Double-stranded cDNA was prepared from cell line RNA. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGCCCATATGACC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCGCGCGCGCATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.65 kb (range 0.9-4.0 kb). 14/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

ORIGIN

Query Match 4.3%; Score 163; DB 10; Length 848;
Best Local Similarity 100.0%; Pred. No. 2.7e-22;
Matches 163; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3049 CATCAATGCTCTTCTGCTGCTTCCACCATAGCCACAGCAGATCTTAAATCTTTCCGAGAT 3108
DB 1 CATCAATGCTCTTCTGCTGCTTCCACCATAGCCACAGCAGATCTTAAATCTTTCCGAGAT 60

QY 3109 CTAGTTTTTTCAGCAAGCAGGATTTAAGAAATGTAATCTTATGTTGTTATGAGAAC 3168
DB 61 CTAGTTTTTTCAGCAAGCAGGATTTAAGAAATGTAATCTTATGTTGTTATGAGAAC 120

QY 3169 AATAGATCATGCTCTATAGTGCTTTTAACTGTAATTT 3211
DB 121 AATAGATCATGCTCTATAGTGCTTTTAACTGTAATTT 163

RESULT 18

AG182348

LOCUS Pan troglodytes DNA, clone: RP43-055101.TJ, genomic survey

DEFINITION sequence.

ACCESSION AG182348

VERSION AG182348.1 GI:16712028

KEYWORDS GSS.

SOURCE Pan troglodytes (chimpanzee)

ORGANISM Pan troglodytes

REFERENCE 1 (bases 1 to 643)

AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.

TITLE BAC end sequences of Library RPCI-43

JOURNAL Unpublished

COMMENT BAC end sequences of Library RPCI-43

REFERENCE 2 (bases 1 to 643)

AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.

TITLE Direct Submission

JOURNAL Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: chimpanzee@sc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/; Tel: 81-45-503-9111, Fax: 81-45-503-9170)

COMMENT Clones are derived from the chimpanzee BAC library RPCI-43. This BAC end was generated during the RED process and may have higher chance of clone tracking errors.

PRIMERS

Sequencing: TJ

LIBRARY

Vector: pBACe3.6

R.site 1: EcorI

R.site 2: EcorI

Location/Qualifiers

1. .643

/organism="Pan troglodytes"

FEATURES

source

Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 1 GAGAAACCGAGTCACGTGTAAGATGCGAAATTATCTCTCGAAAACCTCAG 53
 |||||
 Db 42 GAGAAACCGAGTCACGTGTAAGATGCGAAATTATCTCTCGAAAACCTCAG 94
 |||||

RESULT 20
 CA301835
 LOCUS
 DEFINITION
 CA301835.1 GI:24464894
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Hydra magnipapillata
 Hydra magnipapillata
 Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Anthomedusae;
 Hydridae; Hydra.

REFERENCE
 1 (bases 1 to 344)
 Bode,H., Blumberg,B., Steele,R., Wigge,P., Gee,L., Nguyen,Q.,
 Martinez,D., Kibler,D., Hampson,S., Clifton,S., Pape,D., Marra,M.,
 Hillier,L., Martin,J., Wyllie,T., Dante,M., Theising,B., Bowers,Y.,
 Gibbons,M., Ritter,E., Bennett,J., Ronko,I., Tsagareishvili,R.,
 Maguire,L., Kennedy,S., Waterston,R. and Wilson,R.
 WashU Hydra EST Project
 Unpublished (2002)
 Contact: Hans Bode
 WashU Hydra EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu

TITLE
 JOURNAL
 COMMENT
 Library was constructed by H. Bode and B. Blumberg DNA sequencing
 by: Washington University Genome Sequencing Center For information
 on obtaining a clone please contact: Hans Bode (hnbode@wustl.edu)
 Seq primer: -4ORP from Gibco
 High quality sequence stop: 337.
 Location/Qualifiers
 1..344
 /organism="Hydra magnipapillata"
 /mol_type="mRNA"
 /db_xref="taxon:6085"
 /lab_host="DH5alpha"
 /clone_lib="Hydra cDNA library"
 /notes="Vector: pSPOT6; Site 1: Noti; Site 2: SalI;
 Libraries prepared by Phil Wigge, Plant Molecular and
 Cellular Biology, Lab (Detlev Weigel), The Salk Institute
 for Biological Science, 10010 North Torrey Pines Road, La
 Jolla, CA 92037."

ORIGIN
 Query Match 1.4%; Score 52; DB 14; Length 344;
 Best Local Similarity 100.0%; Pred. No. 0.62;
 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3711 TATAAATGAAAAA 3762
 |||||
 Db 241 TATAAATGAAAAA 292
 |||||

RESULT 21
 AV706915
 LOCUS
 DEFINITION
 AV706915 ADB Homo sapiens cDNA clone ADBXC08 5', mRNA sequence.
 AV706915
 AV706915.1 GI:10724188
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

REFERENCE
1 (bases 1 to 733)
PENG, Y., Song, H., Huang, Q., Huang, C., Gu, Y., Yang, Y., Gao, G.,
Xiao, H., Xu, X., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z.,
Xu, Z., Zeng, L., Lu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S.,
Zhong, M., Lu, G., Hu, R., Chen, J., Chen, Z. and Han, Z.
Homo sapiens cDNA ADB clones
Unpublished (2000)
Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@hgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES
Location/Qualifiers
1..733
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="ADBXCO8"
/tissue_type="Adrenal gland"
/dev_stage="Adult"
/lab_host="SOLR"
/clone_lib="ADB"
/notes="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"

ORIGIN
Query Match 1.4%; Score 52; DB 9; Length 733;
Best Local Similarity 100.0%; Pred. No. 0.29; Indels 0; Gaps 0;
Matches 52; Conservative 0; Mismatches 0;

QY 3711 TATAATGAAAAA 3762
|||||
Db 108 TATAATGAAAAA 159
|||||

RESULT 22
CA780165
LOCUS 190 bp mRNA linear EST 03-DEC-2002
DEFINITION MPL384.4.H01 MPL Sus scrofa cDNA clone pSPORT1 5', mRNA sequence.
ACCESSION CA780165
VERSION CA780165.1 GI:26018151
KEYWORDS EST.
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (bases 1 to 190)
AUTHORS Center for Animal Functional Genomics.
TITLE Generation of ESTs from mixed pig cDNA libraries
JOURNAL Unpublished (2002)
COMMENT Contact: Steven P. Suchyta
Center for Animal Functional Genomics, Department of Animal Science
Michigan State University
B215 Anthony Hall, East Lansing, MI 48824, USA
Tel: 517 355 8443
Fax: 517 432 9168
Email: suchyta@msu.edu
Single Pass sequencing. Bases called and alt-trimmed with phred
v0.0204425.c. Vector identified by cross_match with the -minscore
20 -minmatch 12 options.
Seq primer: T7.
Location/Qualifiers
1..190
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/clone="pSPORT1"
/sex="Male and female"
/tissue_type="pooled"
/dev_stage="pooled"
/lab_host="MPL"
/clone_lib="MPL"
/notes="Organ: pooled; Vector: pSPORT1; Site_1: NotI;
Site_2: SalI; Library made from pooled tissue from
adipose, adrenal gland, blood leukocytes, brain,
cartilage, eye, heart, intestine, kidney, liver, lung
lymph nodes, mammary gland, myogenic satellite cells,
ovary, pancreas, pituitary gland, placenta, skin, spinal
cord, spleen, stomach, tendon, testes, uterus, and
vascular from various developmental and physiological
stages."

FEATURES
Location/Qualifiers
1..190
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/clone="pSPORT1"
/sex="Male and female"
/tissue_type="pooled"
/dev_stage="pooled"
/lab_host="MPL"
/clone_lib="MPL"
/notes="Organ: pooled; Vector: pSPORT1; Site_1: NotI;
Site_2: SalI; Library made from pooled tissue from
adipose, adrenal gland, blood leukocytes, brain,
cartilage, eye, heart, intestine, kidney, liver, lung
lymph nodes, mammary gland, myogenic satellite cells,
ovary, pancreas, pituitary gland, placenta, skin, spinal
cord, spleen, stomach, tendon, testes, uterus, and
vascular from various developmental and physiological
stages."

ORIGIN
Query Match 1.4%; Score 51; DB 14; Length 322;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

/lab_host="DH108"
/clone_lib="MPL"
/notes="Organ: pooled; Vector: pSPORT1; Site_1: NotI;
Site_2: SalI; Library made from pooled tissue from
adipose, adrenal gland, blood leukocytes, brain,
cartilage, eye, heart, intestine, kidney, liver, lung
lymph nodes, mammary gland, myogenic satellite cells,
ovary, pancreas, pituitary gland, placenta, skin, spinal
cord, spleen, stomach, tendon, testes, uterus, and
vascular from various developmental and physiological
stages."

ORIGIN
Query Match 1.4%; Score 51; DB 14; Length 190;
Best Local Similarity 100.0%; Pred. No. 1.7; Indels 0; Gaps 0;
Matches 51; Conservative 0; Mismatches 0;

QY 3712 ATAAATGAAAAA 3762
|||||
Db 92 ATAAATGAAAAA 142
|||||

RESULT 23
CA779205
LOCUS 322 bp mRNA linear EST 03-DEC-2002
DEFINITION MPL384.8.120 MPL Sus scrofa cDNA clone pSPORT1 5', mRNA sequence.
ACCESSION CA779205
VERSION CA779205.1 GI:26017100
KEYWORDS EST.
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (bases 1 to 322)
AUTHORS Center for Animal Functional Genomics.
TITLE Generation of ESTs from mixed pig cDNA libraries
JOURNAL Unpublished (2002)
COMMENT Contact: Steven P. Suchyta
Center for Animal Functional Genomics, Department of Animal Science
Michigan State University
B215 Anthony Hall, East Lansing, MI 48824, USA
Tel: 517 355 8443
Fax: 517 432 9168
Email: suchyta@msu.edu
Single Pass sequencing. Bases called and alt-trimmed with phred
v0.0204425.c. Vector identified by cross_match with the -minscore
20 -minmatch 12 options.
Seq primer: T7.
Location/Qualifiers
1..322
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/clone="pSPORT1"
/sex="Male and female"
/tissue_type="pooled"
/dev_stage="pooled"
/lab_host="DH108"
/clone_lib="MPL"
/notes="Organ: pooled; Vector: pSPORT1; Site_1: NotI;
Site_2: SalI; Library made from pooled tissue from
adipose, adrenal gland, blood leukocytes, brain,
cartilage, eye, heart, intestine, kidney, liver, lung
lymph nodes, mammary gland, myogenic satellite cells,
ovary, pancreas, pituitary gland, placenta, skin, spinal
cord, spleen, stomach, tendon, testes, uterus, and
vascular from various developmental and physiological
stages."

FEATURES
Location/Qualifiers
1..322
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/clone="pSPORT1"
/sex="Male and female"
/tissue_type="pooled"
/dev_stage="pooled"
/lab_host="DH108"
/clone_lib="MPL"
/notes="Organ: pooled; Vector: pSPORT1; Site_1: NotI;
Site_2: SalI; Library made from pooled tissue from
adipose, adrenal gland, blood leukocytes, brain,
cartilage, eye, heart, intestine, kidney, liver, lung
lymph nodes, mammary gland, myogenic satellite cells,
ovary, pancreas, pituitary gland, placenta, skin, spinal
cord, spleen, stomach, tendon, testes, uterus, and
vascular from various developmental and physiological
stages."

ORIGIN
Query Match 1.4%; Score 51; DB 14; Length 322;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

VERSION BX370381.1 GI:30447935

```

KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 931)
JOURNAL Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
COMMENT Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 5501.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CSOBAG044ZE08_CS04189_1&cluster=5501.r.
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CSOBAG044ZE08_CS04189_1.

FEATURES
Location/Qualifiers
1..931
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODK003VN10"
/cell_line="HELA"
/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 1.4%; Score 51; DB 13; Length 931;
Best Local Similarity 100.0%; Pred. No. 0.35; Indels 0; Gaps 0;
Matches 51; Conservative 0; Mismatches 0;

QY 3712 ATAAATGAAAAA 1201 bp mRNA linear EST 08-MAY-2003
||||| 1201 bp mRNA linear EST 08-MAY-2003
||||| Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
Db 154 ATAAATGAAAAA 1201 bp mRNA linear EST 08-MAY-2003
||||| 1201 bp mRNA linear EST 08-MAY-2003
||||| Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens

RESULT 27
BX387156 1201 bp mRNA linear EST 08-MAY-2003
LOCUS BX387156 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
DEFINITION cDNA clone CSODC010YP24 5-PRIME, mRNA sequence.
ACCESSION BX387156.1 GI:30455496
VERSION BX387156.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 1201)
JOURNAL Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
COMMENT Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 3161.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS5AA004ZA04QP1&cluster=3161.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS5AA004ZA04QP1.

```

```

FEATURES
Location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODC010YP24"
/tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 1.4%; Score 51; DB 13; Length 1201;
Best Local Similarity 100.0%; Pred. No. 0.27; Indels 0; Gaps 0;
Matches 51; Conservative 0; Mismatches 0;

QY 3712 ATAAATGAAAAA 80 bp mRNA linear EST 18-AUG-2003
||||| 80 bp mRNA linear EST 18-AUG-2003
||||| sativa cDNA clone NACL--03-M07, mRNA sequence.
Db 703 ATAAATGAAAAA 80 bp mRNA linear EST 18-AUG-2003
||||| 80 bp mRNA linear EST 18-AUG-2003
||||| sativa cDNA clone NACL--03-M07, mRNA sequence.

RESULT 28
CF328728 80 bp mRNA linear EST 18-AUG-2003
LOCUS CF328728 sativa cDNA clone NACL--03-M07, mRNA sequence.
DEFINITION NACL--03-M07.g1 Rice callus plasmid cDNA library (NACL) Oryza
ACCESSION CF328728.1 GI:33905694
VERSION CF328728.1
KEYWORDS EST.
SOURCE Oryza sativa
ORGANISM Oryza sativa
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
TITLE Erihartoideae; Oryzaceae; Oryza.
JOURNAL 1 (bases 1 to 80);
COMMENT Kim, J.S., Jun, X.M., Cheong, P.J., Kim, M.J., Lee, T.H., Shin, Y.C.,
Song, S.I., Kim, J.K., Kim, Y.-K. and Nam, B.H.
Large-scale Sequencing Analysis of Rice ESTs
Unpublished (2003)
Contact: Nahm B.H.
Genomics and Genetics Institute, GreenGene Biotech Inc.; Division
of Bioscience and Bioinformatics, Myongji University
Yongin, Gyeonggi, Korea
Tel: 82 31 330 6193
Fax: 82 31 321 6355
Email: bhnahm@gbio.com, bhnahm@bio.myongji.ac.kr.

FEATURES
Location/Qualifiers
1..80
/organism="Oryza sativa"
/mol_type="mRNA"
/cultivar="Nackdong"
/db_xref="taxon:4530"
/clone="NACL--03-M07"
/tissue_type="callus"
/dev_stage="proliferated callus on 2N6 media for 30 days"
/lab_host="E.coli DH10B"
/clone_lib="Rice callus plasmid cDNA library (NACL)"
/notes="Vector: PCR-TOPO; Site 1: EcoRI; mRNA was capped
with oligoribonucleotides and then used as templates for
RT-PCR."

ORIGIN
Query Match 1.3%; Score 50; DB 14; Length 80;
Best Local Similarity 100.0%; Pred. No. 6.3; Indels 0; Gaps 0;
Matches 50; Conservative 0; Mismatches 0;

QY 3713 TAAATGAAAAA 3762
||||| 3762
||||| TAAATGAAAAA 3762
Db 56 TAAATGAAAAA 3762
||||| 3762
||||| TAAATGAAAAA 3762

```

RESULT 29

CF929023 1af75f02.y1 S1EP Mus musculus cDNA 5', mRNA linear EST 12-NOV-2003
 DEFINITION laf75f02.y1 S1EP Mus musculus cDNA 5', mRNA sequence.

ACCESSION

CF929023
 VERSION CF929023.1 GI:38276827

KEYWORDS

EST

SOURCE

Mus musculus (house mouse)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 142)
 Tidwell, R., Clifton, S., Marra, M., Hillier, L., Pape, D., Martin, J.,
 Wylie, T., Theising, B., Bowers, Y., Gibbons, M., Ritter, E., Bennet, J.,
 Renko, I., Tsagarisvili, R., Belaygorod, L., Grow, A., Maguire, L.,
 Waterston, R. and Wilson, R.

TITLE

WashU Stem cell EST Project

JOURNAL

Unpublished (2002)

COMMENT

Contact: Jeff Gordon and Mike Lovett
 WashU, Human Genetics Division
 Washington University School of Medicine
 Library materials provided by: Stappenbeck Th.S. and Gordon J.I.
 Library constructed by: Korshunova Y. and Lovett M. DNA sequencing
 by: Washington University Genome Sequencing Center For information
 on obtaining a clone please contact: Rose Tidwell
 (rtidwell@genetics.wustl.edu)

Putative full length read

vector to vector length is

Seq primer: -40RP from Gibco.

Location/Qualifiers

FEATURES

source

1..142

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/tissue_type="small intestinal epithelial progenitors"

/lab_host="DMSalpha Ultra Max cells (Invitrogen)"

/clone_lib="S1EP"

/notes="Vector: pAMP1; The library was synthesized with
 modified SMART primers with dUTP at the end. After
 treatment with UDG the cDNA was cloned in pAMP1 vector by
 annealing."

ORIGIN

Query Match 1.3%; Score 50; DB 14; Length 142;
 Best Local Similarity 100.0%; Pred. No. 3.5;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3713 TAAATGAAA 3762

|||||

Db 91 TAAATGAAA 140

|||||

RESULT 30

CF755431 lae60c03.y2 Colon epithelia progenitors cDNA Mus musculus cDNA 5',
 DEFINITION lae60c03.y2 Colon epithelia progenitors cDNA Mus musculus cDNA 5',
 mRNA sequence.

ACCESSION

CF755431

VERSION

CF755431.1 GI:37699236

KEYWORDS

EST

SOURCE

Mus musculus (house mouse)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 146)
 Tidwell, R., Clifton, S., Marra, M., Hillier, L., Pape, D., Martin, J.,
 Wylie, T., Theising, B., Bowers, Y., Gibbons, M., Ritter, E., Bennet, J.,
 Renko, I., Tsagarisvili, R., Belaygorod, L., Grow, A., Maguire, L.,
 Waterston, R. and Wilson, R.

TITLE

WashU Stem cell EST Project

JOURNAL

Unpublished (2002)

COMMENT

Contact: Jeff Gordon and Mike Lovett

WashU, Human Genetics Division

Washington University School of Medicine

1st strand of cDNA was synthesized with reverse transcriptase and
 oligo(dT) beads, then cDNA was amplified by PCR using modified
 SMART primers. The final cDNA was cloned in pAMP1 vector in
 annealing reaction with Uracil DNA Glycosylase (UDG). Library
 constructed by Y.Korshunova and M. Lovett. Library materials
 provided by J. Gordon lab.

Putative full length read

vector to vector length is

Seq primer: -40RP from Gibco.

Location/Qualifiers

FEATURES

source

1..146

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/tissue_type="colon epithelia progenitor cells"

/lab_host="DMSalpha"

/clone_lib="Colon epithelia progenitors cDNA"

/note="Vector: pAMP1; 1st strand of cDNA was synthesized
 with reverse transcriptase and oligo(dT) beads, then cDNA
 was amplified by PCR using modified SMART primers. The
 final cDNA was cloned in pAMP1 vector in annealing
 reaction with Uracil DNA Glycosylase (UDG). Library
 constructed by Y.Korshunova and M. Lovett. Library
 materials provided by J. Gordon lab."

ORIGIN

Query Match 1.3%; Score 50; DB 14; Length 146;
 Best Local Similarity 100.0%; Pred. No. 3.4;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3713 TAAATGAAA 3762

|||||

Db 97 TAAATGAAA 146

|||||

RESULT 31

CD645057 lab18e09.y1 Normalized Small intestinal epithelia progenitors Mus
 LOCUS musculus cDNA 5', mRNA sequence.
 DEFINITION

musculus cDNA 5', mRNA sequence.

CD645057

CD645057.1 GI:31817557

EST.

Mus musculus (house mouse)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 148)

Tidwell, R., Clifton, S., Marra, M., Hillier, L., Pape, D., Martin, J.,
 Wylie, T., Theising, B., Bowers, Y., Gibbons, M., Ritter, E., Bennet, J.,
 Renko, I., Tsagarisvili, R., Belaygorod, L., Grow, A., Maguire, L.,
 Waterston, R. and Wilson, R.

WashU Stem cell EST Project

Unpublished (2002)

Contact: Rose Tidwell

WashU Stem cell EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

Original lib/1st strand of cDNA was synthesized with reverse

transcriptase and oligo(dT) beads, then cDNA was amplified by PCR

using modified SMART primers. The final cDNA was cloned in pAMP1

vector in annealing reaction with Uracil DNA Glycosylase (UDG).

Normalized by generating biotin labeled linear PCR using original

lib cut with Sal I as template, followed with Sau3A digest, and

hybridizing with ssDNA from original lib. RNA provided by Jeff

Gordon lab and library construction by Rose Tidwell and Mike

Lovett.

Putative full length read

vector to vector length is


```

FEATURES
  source
    Seq primer: -40RP from Gibco.
    Location/Qualifiers
      1..148
        /organism="Mus musculus"
        /mol_type="mRNA"
        /db_xref="taxon:10090"
        /tissue_type="small intestinal epithelial progenitor cells"
        /lab_host="X11- Blue MRF"
        /clone_lib="Normalized Small intestinal epithelia
        progenitors"
        /note="vector: pAMP1; Site 1: UDG cloning; Site 2: UDG
        cloning; Original lib: 1st strand of cDNA was synthesized
        with reverse transcriptase and oligo(dT) beads, then cDNA
        was amplified by PCR using modified SMART primers. The
        final cDNA was cloned in pAMP1 vector in annealing
        reaction with Uracil DNA Glycosylase (UDG). Normalized by
        generating Biotin labeled linear PCR using original lib
        cut with Sal I as template, followed with SauaI digest,
        and hybridizing with ssDNA from original lib. RNA provided
        by Jeff Gordon lab and library construction by Rose
        Tidwell and Mike Lovett."

ORIGIN
  Query Match      1.3%; Score 50; DB 14; Length 148;
  Best Local Similarity 100.0%; Pred. No. 3.4;
  Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3713 TAAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
      |||||||
Db 94 TAAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 143

RESULT 32
CF259903
LOCUS
DEFINITION
  lab41h07.y1 Normalized colon epithelial progenitors cDNA library
  Mus musculus cDNA 5', mRNA sequence.
  EST.
ACCESSION
  CF259903.1 GI:33587476
VERSION
  EST.
KEYWORDS
  Mus musculus (house mouse)
SOURCE
  Mus musculus
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
  1 (bases 1 to 148)
  Tidwell,R., Clifton,S., Marra,M., Hillier,L., Pape,D., Martin,J.,
  Wyllie,T., Theising,B., Bowers,Y., Gibbons,M., Ritter,E., Bennet,J.,
  Ronko,I., Tsagarisvilli,R., Belaygorod,L., Grow,A., Maguire,L.,
  Waterston,R. and Wilson,R.
  WashU Stem cell EST Project
  Unpublished (2002)
  Contact: Jeff Gordon and Mike Lovett
  WashU, Human Genetics Division
  Washington University School of Medicine
  1st strand of cDNA was synthesized with reverse transcriptase and
  oligo(dT) beads, then cDNA was amplified by PCR using modified
  SMART primers. The final cDNA was cloned in pAMP1 vector in
  annealing reaction with Uracil DNA Glycosylase (UDG). Library
  materials provided by T.S. Stappenbeck and J.I. Gordon. Library
  constructed by R. Tidwell Y.Korshunova, M. Lovett. Library was
  re-arrayed by R. Tidwell.
  Putative full length read
  vector to vector length is
  Seq primer: -40RP from Gibco.
  Location/Qualifiers
    1..148
      /organism="Mus musculus"
      /mol_type="mRNA"
      /db_xref="taxon:10090"
      /tissue_type="laser-captured colonic epithelial
      progenitors cells from conventionally raised C57/Bl6 mice"
      /lab_host="X11 Blue MRF"
      /clone_lib="Normalized colon epithelial progenitors cDNA

FEATURES
  source
    Seq primer: -40RP from Gibco.
    Location/Qualifiers
      1..148
        /organism="Mus musculus"
        /mol_type="mRNA"
        /db_xref="taxon:10090"
        /tissue_type="laser-captured colonic epithelial
        progenitors cells from conventionally raised C57/Bl6 mice"
        /lab_host="X11 Blue MRF"
        /clone_lib="Normalized colon epithelial progenitors cDNA

```

```

library"
/note="vector: pAMP1; 1st strand of cDNA was synthesized
with reverse transcriptase and oligo(dT) beads, then cDNA
was amplified by PCR using modified SMART primers. The
final cDNA was cloned in pAMP1 vector in annealing
reaction with Uracil DNA Glycosylase (UDG). Library
materials provided by T.S. Stappenbeck and J.I. Gordon.
Library constructed by R. Tidwell Y.Korshunova, M. Lovett.
Library was re-arrayed by R. Tidwell."

ORIGIN
  Query Match      1.3%; Score 50; DB 14; Length 148;
  Best Local Similarity 100.0%; Pred. No. 3.4;
  Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3713 TAAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
      |||||||
Db 97 TAAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 146

RESULT 33
CF545723
LOCUS
DEFINITION
  lad99f06.y1 Colon epithelia progenitors cDNA Mus musculus cDNA 5',
  mRNA sequence.
  EST.
ACCESSION
  CF545723.1 GI:34893153
VERSION
  EST.
KEYWORDS
  Mus musculus (house mouse)
SOURCE
  Mus musculus
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
  1 (bases 1 to 149)
  Tidwell,R., Clifton,S., Marra,M., Hillier,L., Pape,D., Martin,J.,
  Wyllie,T., Theising,B., Bowers,Y., Gibbons,M., Ritter,E., Bennet,J.,
  Ronko,I., Tsagarisvilli,R., Belaygorod,L., Grow,A., Maguire,L.,
  Waterston,R. and Wilson,R.
  WashU Stem cell EST Project
  Unpublished (2002)
  Contact: Jeff Gordon and Mike Lovett
  WashU, Human Genetics Division
  Washington University School of Medicine
  1st strand of cDNA was synthesized with reverse transcriptase and
  oligo(dT) beads, then cDNA was amplified by PCR using modified
  SMART primers. The final cDNA was cloned in pAMP1 vector in
  annealing reaction with Uracil DNA Glycosylase (UDG). Library
  constructed by Y.Korshunova and M. Lovett. Library materials
  provided by J. Gordon lab.
  Putative full length read
  vector to vector length is
  Seq primer: -40RP from Gibco.
  Location/Qualifiers
    1..149
      /organism="Mus musculus"
      /mol_type="mRNA"
      /db_xref="taxon:10090"
      /tissue_type="colon epithelia progenitor cells"
      /lab_host="DH5alpha"
      /clone_lib="Colon epithelia progenitors cDNA"
      /note="vector: pAMP1; 1st strand of cDNA was synthesized
      with reverse transcriptase and oligo(dT) beads, then cDNA
      was amplified by PCR using modified SMART primers. The
      final cDNA was cloned in pAMP1 vector in annealing
      reaction with Uracil DNA Glycosylase (UDG). Library
      constructed by Y.Korshunova and M. Lovett. Library
      materials provided by J. Gordon lab."

FEATURES
  source
    Seq primer: -40RP from Gibco.
    Location/Qualifiers
      1..149
        /organism="Mus musculus"
        /mol_type="mRNA"
        /db_xref="taxon:10090"
        /tissue_type="colon epithelia progenitor cells"
        /lab_host="DH5alpha"
        /clone_lib="Colon epithelia progenitors cDNA"
        /note="vector: pAMP1; 1st strand of cDNA was synthesized
        with reverse transcriptase and oligo(dT) beads, then cDNA
        was amplified by PCR using modified SMART primers. The
        final cDNA was cloned in pAMP1 vector in annealing
        reaction with Uracil DNA Glycosylase (UDG). Library
        constructed by Y.Korshunova and M. Lovett. Library
        materials provided by J. Gordon lab."

ORIGIN
  Query Match      1.3%; Score 50; DB 14; Length 149;
  Best Local Similarity 100.0%; Pred. No. 3.4;
  Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3713 TAAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762

```

TITLE

modified SMART primers with dUTP at the end. After treatment with UDG the cDNA was cloned in pAMP1 vector by annealing."

ORIGIN

Query Match 1.3%; Score 50; DB 14; Length 158;
Best Local Similarity 100.0%; Pred. No. 3.2;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3713 TAAATGAAAAA... 3762
Db 94 TAAATGAAAAA... 143

RESULT 37

AW315672 171 bp mRNA linear EST 09-JUL-2000
LOCUS 13217 MARC 2P1G Sus scrofa cDNA 5', mRNA sequence.
DEFINITION AW315672

VERSION AW315672.1 GI:6744928

KEYWORDS

SOURCE Sus scrofa (pig)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

REFERENCE

AUTHORS Fahnenkrug, S.C., Smith, T.P.L., Freking, B.A., Cho, J., White, J.,
Vallet, J., Wise, T., Rohrer, G.A., Perte, G., Sultana, R.,
Quackenbush, J. and Keele, J.W.

Porcine gene discovery by normalized cDNA-library sequencing and

TITLE

EST cluster assembly

JOURNAL

MEDLINE Mamm. Genome 13 (8), 475-478 (2002)

PUBMED

22213789

COMMENT

Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 20
and -minmatch 12 options.

PCR Primers

FORWARD: AGGAACAGCTATGACCAT

BACKWARD: GTTTCAGTCAGCAGC

Plate: 9 row: B column: 12

Seq primer: ATTAGTGACACTATAG.

Location/Qualifiers

FEATURES

1..171

/organism="Sus scrofa"

/mol_type="mRNA"

/db_xref="taxon:9823"

/tissue_type="pooled"

/lab_host="DH10B"

/clone_lib="MARC 2P1G"

/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;

Library made from pooled tissue from testis, ovary,

endometrium, hypothalamus, pituitary, and placenta."

ORIGIN

Query Match 1.3%; Score 50; DB 10; Length 171;
Best Local Similarity 100.0%; Pred. No. 2.9;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3713 TAAATGAAAAA... 3762
Db 113 TAAATGAAAAA... 162

RESULT 38

CF354150 174 bp mRNA linear EST 20-AUG-2003
LOCUS

DEFINITION

CF354150

VERSION CF354150.1

KEYWORDS

SOURCE

ORGANISM

Mus musculus (house mouse)

REFERENCE

AUTHORS

1 (bases 1 to 174)

Tidwell, R., Clifton, S., Marra, M., Hillier, L., Pape, D., Martin, J.,

Wyllie, I., Theising, B., Bowers, Y., Gibbons, M., Ritter, E., Bennet, J.,

Ronko, I., Tsagaris, R., Belaygorod, L., Grow, A., Maguire, L.,

Waterston, R. and Wilson, R.

WashU Stem Cell EST Project

Unpublished (2002)

Contact: Jeff Gordon and Mike Lovett

WashU, Human Genetics Division

Washington University School of Medicine

Library materials provided by: Stappenbeck Th.S. and Gordon J.I.

Library constructed by: Korshunova Y. and Lovett M. DNA sequencing

by: Washington University Genome Sequencing Center For information

on obtaining a clone please contact: Rose Tidwell

(rtidwell@genetics.wustl.edu)

Putative full length read

vector to vector length is

Seq primer: -40RP from Gibco.

Location/Qualifiers

1..174

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/tissue_type="small intestinal epithelial progenitors"

/lab_host="DH5alpha Ultra Max cells (Invitrogen)"

/clone_lib="SiEP"

/note="Vector: PAMP1; The library was synthesized with

modified SMART primers with dUTP at the end. After

treatment with UDG the cDNA was cloned in pAMP1 vector by

annealing."

Query Match 1.3%; Score 50; DB 14; Length 174;

Best Local Similarity 100.0%; Pred. No. 2.9;

Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3713 TAAATGAAAAA... 3762

Db 97 TAAATGAAAAA... 146

RESULT 39

BM177915

LOCUS

DEFINITION

saj67a02.y1 Gm-c1072 Glycine max cDNA clone SOYBEAN CLONE ID:

Gm-c1072-4612 5', mRNA sequence.

BM177915

VERSION

BM177915.1 GI:117401133

KEYWORDS

SOURCE

ORGANISM

Glycine max (soybean)

Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;

Glycine.

1 (bases 1 to 175)

Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Coryell, V.,

Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,

Back, C., Wyllie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M.,

Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N.,

Schuck, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,

McCann, R., Waterston, R. and Wilson, R.

Public Soybean EST Project

Unpublished (1999)

Contact: Shoemaker R/Public Soybean EST Project

TITLE

JOURNAL

COMMENT

Waterston, R. and Wilson, R.
WashU Stem cell EST Project
Unpublished (2002)
Contact: Jeff Gordon and Mike Lovett
WashU, Human Genetics Division
Washington University School of Medicine
Library materials provided by: Stappenbeck Th.S. and Gordon J.I.
Library constructed by: Korshunova Y. and Lovett M. DNA sequencing
by: Washington University Genome Sequencing Center For information
on obtaining a clone please contact: Rose Tidwell
(rtidwell@genetics.wustl.edu)
Putative full length read
vector to vector length is
seg primer: -40RP from Gibco.

```

FEATURES
source
Seq primer: -40RP from Gibco.
Location/Qualifiers
1, 183
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/tissue_type="small intestine"
/lab_host="DH5alpha Ultr"
/clone_lib="SIEP"
/note="Vector: pAMP1; Th
modified SMART primers w
treatment with UDG the c
appealing."

```

```

Query Match      1.3%; Score 50; DB 14; Length 183;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
db TAAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 143

```

REFERENCE
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus
1 (Bases 1 to 183)
AUTHORS
Tidwell, R., Clifton, S., Mowers, M., Hillier, L., Pape, D., Martin, J.,
Wyell, T., Theising, B., Bowers, G., Gibbons, M., Ritter, E., Bannet, J.,
Ronko, I., Teagareishvili, R., Belaygorod, L., Grow, A., Maguire, L.,
Waterson, R. and Wilson, R.
TITLE
WashU Stem cell EST Project
Unpublished (2002)

WashU, Human Genetics Division
Washington University School of Medicine
1st strand of cDNA was synthesized with reverse transcriptase and
oligo(dt) beads, then cDNA was amplified by PCR using modified
SMART primers. The final cDNA was cloned in pAMP1 vector in
annealing reaction with uracil DNA Glycosylase (UDG). Library
constructed by Y. Korshunova and M. Lovett. Library materials

```

Putative full length read
vector to vector length is
Seq primer: -40Rp from Gibco.
Location/Qualifiers
1..183
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
FEATURES
source

```

```

/tissue type="colon epithelia progenitor cells"
/lab host="DH5alpha"
/clone_lib="Colon epithelia progenitors cDNA"
/notes="vector: pAMP1; 1st strand of cDNA was synthesized
with reverse transcriptase and oligo(dT) beads, then cDNA
was amplified by PCR using modified SMART primers. The
final cDNA was cloned in pAMP1 vector in annealing
reaction with Uracil DNA Glycosylase (UDG). Library
constructed by Y. Korshunova and M. Lovett. Library
materials provided by J. Gordon lab."

ORIGIN
Query Match          1.3%; Score 50; DB 14; Length 183;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3713 TAAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 94 TAAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 143

RESULT 42
CF970587 lag77f05.y1 Colon epithelia progenitors cDNA Mus musculus cDNA 5',
LOCUS mRNA sequence.
DEFINITION lag77f05.y1 Colon epithelia progenitors cDNA Mus musculus cDNA 5',
ACCESSION CF970587
VERSION CF970587
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 208)
AUTHORS Tidwell,R., Clifton,S., Marra,M., Hillier,L., Pape,D., Martin,J.,
Wyllie,T., Theising,B., Bowers,Y., Gibbons,M., Ritter,E., Bennet,J.,
Ronko,I., Tsagarisvilli,R., Belaygorod,L., Grow,A., Maguire,L.,
Waterson,R. and Wilson,R.
WashU Stem cell EST Project
TITLE Unpublished (2002)
JOURNAL
COMMENT Contact: Jeff Gordon and Mike Lovett
Washington University School of Medicine
1st strand of cDNA was synthesized with reverse transcriptase and
oligo(dT) beads, then cDNA was amplified by PCR using modified
SMART primers. The final cDNA was cloned in pAMP1 vector in
annealing reaction with Uracil DNA Glycosylase (UDG). Library
constructed by Y. Korshunova and M. Lovett. Library materials
provided by J. Gordon lab.
Putative full length read
vector to vector length is
Seq primer: -40RP from Gibco
High quality sequence stop: 208.
Location/Qualifiers
1. .208
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/tissue type="colon epithelia progenitor cells"
/clone_lib="DH5alpha"
/lab host="Colon epithelia progenitors cDNA"
/notes="vector: pAMP1; 1st strand of cDNA was synthesized
with reverse transcriptase and oligo(dT) beads, then cDNA
was amplified by PCR using modified SMART primers. The
final cDNA was cloned in pAMP1 vector in annealing
reaction with Uracil DNA Glycosylase (UDG). Library
constructed by Y. Korshunova and M. Lovett. Library
materials provided by J. Gordon lab."

ORIGIN
Query Match          1.3%; Score 50; DB 14; Length 208;
Best Local Similarity 100.0%; Pred. No. 2.4;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3713 TAAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 94 TAAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 143

RESULT 44
CF352096 lab55c12.y1 Colon epithelia progenitors cDNA Mus musculus cDNA 5',
LOCUS mRNA sequence.
DEFINITION lab55c12.y1 Colon epithelia progenitors cDNA Mus musculus cDNA 5',
ACCESSION CF352096
VERSION CF352096
KEYWORDS EST.
SOURCE Mus musculus (house mouse)

```

```

Qy 3713 TAAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 94 TAAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 143

RESULT 43
CF425041 ladi4d09.y1 Colon epithelia progenitors cDNA Mus musculus cDNA 5',
LOCUS mRNA sequence.
DEFINITION ladi4d09.y1 Colon epithelia progenitors cDNA Mus musculus cDNA 5',
ACCESSION CF425041
VERSION CF425041
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 215)
AUTHORS Tidwell,R., Clifton,S., Marra,M., Hillier,L., Pape,D., Martin,J.,
Wyllie,T., Theising,B., Bowers,Y., Gibbons,M., Ritter,E., Bennet,J.,
Ronko,I., Tsagarisvilli,R., Belaygorod,L., Grow,A., Maguire,L.,
Waterson,R. and Wilson,R.
WashU Stem cell EST Project
TITLE Unpublished (2002)
JOURNAL
COMMENT Contact: Jeff Gordon and Mike Lovett
Washington University School of Medicine
1st strand of cDNA was synthesized with reverse transcriptase and
oligo(dT) beads, then cDNA was amplified by PCR using modified
SMART primers. The final cDNA was cloned in pAMP1 vector in
annealing reaction with Uracil DNA Glycosylase (UDG). Library
constructed by Y. Korshunova and M. Lovett. Library materials
provided by J. Gordon lab.
Putative full length read
vector to vector length is
Seq primer: -40RP from Gibco
High quality sequence stop: 215.
Location/Qualifiers
1. .215
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/tissue type="colon epithelia progenitor cells"
/clone_lib="DH5alpha"
/lab host="Colon epithelia progenitors cDNA"
/notes="vector: pAMP1; 1st strand of cDNA was synthesized
with reverse transcriptase and oligo(dT) beads, then cDNA
was amplified by PCR using modified SMART primers. The
final cDNA was cloned in pAMP1 vector in annealing
reaction with Uracil DNA Glycosylase (UDG). Library
constructed by Y. Korshunova and M. Lovett. Library
materials provided by J. Gordon lab."

ORIGIN
Query Match          1.3%; Score 50; DB 14; Length 215;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3713 TAAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 94 TAAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 143

RESULT 44
CF352096 lab55c12.y1 Colon epithelia progenitors cDNA Mus musculus cDNA 5',
LOCUS mRNA sequence.
DEFINITION lab55c12.y1 Colon epithelia progenitors cDNA Mus musculus cDNA 5',
ACCESSION CF352096
VERSION CF352096
KEYWORDS EST.
SOURCE Mus musculus (house mouse)

```

ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 230)
AUTHORS Tidwell,R., Clifton,S., Marra,M., Hillier,L., Pape,D., Martin,J., Wylie,T., Theising,B., Bowers,Y., Gibbons,M., Ritter,E., Bennet,J., Ronko,I., Tsagarisshvili,R., Belaygorod,L., Grow,A., Maguire,L., Waterston,R. and Wilson,R.
TITLE WashU Stem cell EST Project
JOURNAL Unpublished (2002)
COMMENT Contact: Jeff Gordon and Mike Lovett
Washington University School of Medicine
1st strand of cDNA was synthesized with reverse transcriptase and oligo(dT) beads, then cDNA was amplified by PCR using modified SMART primers. The final cDNA was cloned in pAMP1 vector in annealing reaction with Uracil DNA Glycosylase (UDG). Library constructed by Y.Korshunova and M. Lovett. Library materials provided by J. Gordon lab.
Putative full length read
vector to vector length is
Seq primer: -40RP from Gibco.
Location/Qualifiers
1. .230
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/tissue_type="colon epithelia progenitor cells"
/lab_host="DH5alpha"
/clone_lib="Colon epithelia progenitors cDNA"
/note="Vector: pAMP1; 1st strand of cDNA was synthesized with reverse transcriptase and oligo(dT) beads, then cDNA was amplified by PCR using modified SMART primers. The final cDNA was cloned in pAMP1 vector in annealing reaction with Uracil DNA Glycosylase (UDG). Library constructed by Y.Korshunova and M. Lovett. Library materials provided by J. Gordon lab."

Query Match 1.3%; Score 50; DB 14; Length 230;
Best Local Similarity 100.0%; Pred.No.2.2; Indels 0; Gaps 0;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3713 TAAATGAAAAA 3762
DB 94 TAAATGAAAAA 143

RESULT 45
AL045558
LOCUS DKFZP434L035.1 233 bp mRNA linear EST 04-SEP-2003
DEFINITION DKFZP434L035.1 434 (synonym: htes3) Homo sapiens cDNA clone
ACCESSION AL045558
VERSION AL045558.2 GI:5936028
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 233)
AUTHORS Duesterhoeft,A., Lauber,J., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
TITLE EST (Duesterhoeft, et al.)
JOURNAL Unpublished (1999)
COMMENT On Jul 9, 1999 this sequence version replaced gi:5433688.
Contact: MIPS
MIPS Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ), Email s.wiemann@dkfz-heidelberg.de; sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing

consortium of the German Genome Project.

No si sequence available.
This clone (DKFZP434L035) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin- Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES

source
1. .233
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZP434L035"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="434 (synonym: htes3)"
/note="Vector: pSPort1; Site_1: NotI; Site_2: SalI"

ORIGIN

Query Match 1.3%; Score 50; DB 9; Length 233;
Best Local Similarity 100.0%; Pred.No.2.2; Indels 0; Gaps 0;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3713 TAAATGAAAAA 3762
DB 183 TAAATGAAAAA 232

RESULT 46

CF926508 249 bp mRNA linear EST 12-NOV-2003
LOCUS lag252908.y1 Colon epithelia progenitors cDNA Mus musculus cDNA 5',
DEFINITION mRNA sequence.
ACCESSION CF926508
VERSION CF926508.1 GI:38271752
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 249)
AUTHORS Tidwell,R., Clifton,S., Marra,M., Hillier,L., Pape,D., Martin,J., Wylie,T., Theising,B., Bowers,Y., Gibbons,M., Ritter,E., Bennet,J., Ronko,I., Tsagarisshvili,R., Belaygorod,L., Grow,A., Maguire,L., Waterston,R. and Wilson,R.
TITLE WashU Stem cell EST Project
JOURNAL Unpublished (2002)
COMMENT Contact: Jeff Gordon and Mike Lovett
Washington University School of Medicine
1st strand of cDNA was synthesized with reverse transcriptase and oligo(dT) beads, then cDNA was amplified by PCR using modified SMART primers. The final cDNA was cloned in pAMP1 vector in annealing reaction with Uracil DNA Glycosylase (UDG). Library constructed by Y.Korshunova and M. Lovett. Library materials provided by J. Gordon lab.
Putative full length read
vector to vector length is
Seq primer: -40RP from Gibco
High quality sequence stop: 249.
Location/Qualifiers
1. .249
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/tissue_type="colon epithelia progenitor cells"
/clone_lib="DH5alpha"
/note="Vector: pAMP1; 1st strand of cDNA was synthesized with reverse transcriptase and oligo(dT) beads, then cDNA was amplified by PCR using modified SMART primers. The final cDNA was cloned in pAMP1 vector in annealing reaction with Uracil DNA Glycosylase (UDG). Library constructed by Y.Korshunova and M. Lovett. Library materials provided by J. Gordon lab."

FEATURES

source
1. .249
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/tissue_type="colon epithelia progenitor cells"
/lab_host="DH5alpha"
/clone_lib="Colon epithelia progenitors cDNA"
/note="Vector: pAMP1; 1st strand of cDNA was synthesized with reverse transcriptase and oligo(dT) beads, then cDNA was amplified by PCR using modified SMART primers. The final cDNA was cloned in pAMP1 vector in annealing reaction with Uracil DNA Glycosylase (UDG). Library constructed by Y.Korshunova and M. Lovett. Library

materials provided by J. Gordon lab."

```

ORIGIN
  Query Match      1.3%; Score 50; DB 14; Length 249;
  Best Local Similarity 100.0%; Pred. No. 2;
  Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3713 TAAATGAAAAA.....AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
      |||||||
Db 94 TAAATGAAAAA.....AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 143

RESULT 47
CF805201
LOCUS
DEFINITION
  CF805201 250 bp mRNA linear EST 24-OCT-2003
  lad64c06.y1 Gastric Epithelial Progenitor 2 Mus musculus cDNA 5',
  mRNA sequence.
ACCESSION
  CF805201
VERSION
  CF805201.1 GI:37974603
KEYWORDS
  EST.
SOURCE
  Mus musculus (house mouse)
ORGANISM
  Mus musculus
REFERENCE
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS
  Tidwell,R., Clifton,S., Marra,M., Hillier,L., Pape,D., Martin,J.,
  Wylie,T., Theising,B., Bowers,Y., Gibbons,M., Ritter,E., Bennett,J.,
  Ronko,I., Tsagarisvili,R., Belaygorod,L., Grow,A., Maguire,L.,
  Waterston,R. and Wilson,R.
TITLE
  WashU Stem cell EST Project
JOURNAL
  Unpublished (2002)
COMMENT
  Contact: Jeff Gordon and Mike Lovett
  WashU, Human Genetics Division
  Washington University School of Medicine
  1st strand of cDNA was synthesized with reverse transcriptase and
  oligo(dT) beads, then cDNA was amplified by PCR using modified
  SMART primers. The final cDNA was cloned in pAMP1 vector in
  annealing reaction with Uracil DNA Glycosylase (UDG). Library
  constructed by Y.Korshunova and M. Lovett. Library materials
  provided by Mills JC & Gordon JI.
  Putative full length read
  vector to vector length is
  Seq primer: -40RP from Gibco.
FEATURES
  source
    1..250
    /organism="Mus musculus"
    /mol_type="mRNA"
    /db_xref="taxon:10090"
    /tissue_type="Laser-captured isthmal cells from tox176
    transgenic mice"
    /dev_stage="adult"
    /lab_host="DH5alpha"
    /clone_lib="Gastric Epithelial Progenitor 2"
    /note="Vector: pAMP1; 1st strand of cDNA was synthesized
    with reverse transcriptase and oligo(dT) beads, then cDNA
    was amplified by PCR using modified SMART primers. The
    final cDNA was cloned in pAMP1 vector in annealing
    reaction with Uracil DNA Glycosylase (UDG). Library
    constructed by Y.Korshunova and M. Lovett. Library
    materials provided by Mills JC & Gordon JI."
ORIGIN
  Query Match      1.3%; Score 50; DB 14; Length 250;
  Best Local Similarity 100.0%; Pred. No. 2;
  Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3713 TAAATGAAAAA.....AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
      |||||||
Db 196 TAAATGAAAAA.....AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 245

RESULT 48
AW022602
LOCUS
DEFINITION
  df41b12.y1 Morton Fetal Cochlea Homo sapiens cDNA clone
  IMAGE:2485966 5', mRNA sequence.
ACCESSION
  AW022602
VERSION
  AW022602.1 GI:5876132
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
ORGANISM
  Homo sapiens
REFERENCE
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS
  1 (bases 1 to 255)
  Robertson,N.G., Khetarpal,U., Guierrez-Espeleta,G.A., Bieber,F.R.
  and Morton,C.C.
TITLE
  Isolation of novel and known genes from a human fetal cochlear cDNA
  library using subtractive hybridization and differential screening
  Genomics 23, 42-50 (1994)
JOURNAL
  95130111
MEDLINE
  7829101
PUBMED
  7829101
COMMENT
  Contact: Morton, C. C.
  Departments of Pathology and Obstetrics, Gynecology and
  Reproductive Biology
  Brigham and Women's Hospital
  75 Francis Street, Harvard Medical School, Boston, MA 02115, USA
  Tel: 617 732 7980
  Fax: 617 738 6996
  Email: cmorton@rics.bwh.harvard.edu
  DNA sequencing and analyses were performed by National Institutes
  of Health Intramural Sequencing Center (NISC; see
  http://www.nisc.nih.gov).
  This clone is available royalty-free through LNL; contact the
  IMAGE Consortium (info@image.llnl.gov) for further information.
  Plate: LLAM6186 row: C column: 23
  Seq primer: M13RPI reverse primer (ABI).
FEATURES
  Location/Qualifiers
    1..255
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="IMAGE:2485966"
    /tissue_type="cochlea"
    /dev_stage="16-22 week fetus"
    /lab_host="SOLR cells (kanamycin resistant)"
    /clone_lib="Morton Fetal Cochlea"
    /note="Organ: ear; Vector: pBluescript SK-; Site 1: EcoRI;
    Site 2: XhoI; Reference: Genomics 23, 42-50 (1994) Cloned
    unidirectionally. Primer: Oligo dT. Fetal cochlea, normal.
    37% of inserts <0.5 kb, 56% 0.5-1.0 kb, 7% >1 kb. Uni-ZAP
    XR Vector. Library constructed by N. Robertson, C. Morton.
    ~5' adaptor sequence: 5' GAATTCGGCAG 3' ~3' adaptor
    sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'"
ORIGIN
  Query Match      1.3%; Score 50; DB 9; Length 255;
  Best Local Similarity 100.0%; Pred. No. 2;
  Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3713 TAAATGAAAAA.....AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
      |||||||
Db 195 TAAATGAAAAA.....AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 244

RESULT 49
BE332081
LOCUS
DEFINITION
  BE332081 259 bp mRNA linear EST 14-JUL-2000
  uc07g01.y1 Soares_thymus_2NDMT Mus musculus cDNA clone
  IMAGE:3327216 5', mRNA sequence.
ACCESSION
  BE332081
VERSION
  BE332081.1 GI:9205857
KEYWORDS
  EST.
SOURCE
  Mus musculus (house mouse)
ORGANISM
  Mus musculus
REFERENCE
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```


hybridizing with ssDNA from original lib. RNA provided by Jeff Gordon lab and library construction by Rose Tidwell and Mike Lovett.

Putative full length read
vector to vector length is
Seq primer: -40RP from Gibco
High quality sequence stop: 261.
Location/Qualifiers

FEATURES

source

```
1. .261
/organism="Mus musculus"
/db_xref="taxon:10090"
/tissue_type="small intestinal epithelial progenitor cells"
/lab_host="XLI- Blue MRF"
/clone_libs="Normalized Small Intestinal Epithelia
progenitors"
/note="Vector: pAMP1; Site 1: UDG cloning; Site 2: UDG
cloning; Original lib: 1st strand of cDNA was synthesized
with reverse transcriptase and oligo(dT) beads, then cDNA
was amplified by PCR using modified SMART primers. The
final cDNA was cloned in pAMP1 vector in annealing
reaction with uracil DNA Glycosylase (UDG). Normalized by
generating biotin labeled linear PCR using original lib
cut with Sal I as template, followed with Sau3A digest,
and hybridizing with ssDNA from original lib. RNA provided
by Jeff Gordon lab and library construction by Rose
Tidwell and Mike Lovett."
```

ORIGIN

Query Match 1.3%; Score 50; DB 14; Length 261;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAAAA 3762
|||||
DB 94 TAAATGAAAAA 143

RESULT 52

AW827290

LOCUS AW827290 265 bp mRNA linear EST 18-MAY-2000
DEFINITION xnl1e09.y1 NCI_CGAP_L15 Homo sapiens cDNA clone IMAGE:2693416 5',
mRNA sequence.

ACCESSION AW827290

VERSION AW827290.1 GI:7921064

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

AUTHORS

TITLE NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

Tumor Gene Index

Unpublished (1997)

JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
R. Emmert-Buck, M.D., Ph.D.
cDNA library Preparation: Life Technologies, Inc.
cDNA library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone Distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL, send email to:
infoimage.llnl.gov

Putative full length read

The vector to vector length is 337

Seq primer: -40RP from Gibco.

Location/Qualifiers

FEATURES

source

```
1. .265
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
```

```
/clone="IMAGE:2693416"
/tissue_type="hepatic adenoma"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_L15"
/note="Organ: liver; Vector: pCMV-SPORT4; Site 1: SalI;  
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 0.8 kb."
```

ORIGIN

Query Match 1.3%; Score 50; DB 10; Length 265;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAAAA 3762
|||||
DB 187 TAAATGAAAAA 236

RESULT 53

CF352983

LOCUS CF352983 269 bp mRNA linear EST 20-AUG-2003
DEFINITION lab60b08.y1 Colon epithelia progenitors cDNA Mus musculus cDNA 5',
mRNA sequence.

ACCESSION CF352983

VERSION CF352983.1 GI:33957496

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 269)
Tidwell,R., Clifton,S., Marra,M., Hillier,L., Pape,D., Martin,J.,
Wyllie,T., Theising,B., Bowers,Y., Gibbons,M., Ritter,E., Bennett,J.,
Ronko,I., Tsagareishvili,R., Belaygorod,L., Grow,A., Maguire,L.,
Waterston,R. and Wilson,R.
WashU Stem cell EST Project

TITLE

Unpublished (2002)

JOURNAL

COMMENT

Contact: Jeff Gordon and Mike Lovett
Washington University School of Medicine
1st strand of cDNA was synthesized with reverse transcriptase and
oligo(dT) beads, then cDNA was amplified by PCR using modified
SMART primers. The final cDNA was cloned in pAMP1 vector in
annealing reaction with uracil DNA Glycosylase (UDG). Library
constructed by Y.Korshunova and M. Lovett. Library materials
provided by J. Gordon lab.

Putative full length read

vector to vector length is

Seq primer: -40RP from Gibco

High quality sequence stop: 269.

Location/Qualifiers

FEATURES

source

```
1. .269
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/tissue_type="colon epithelia progenitor cells"
/lab_host="DH5alpha"
/clone_lib="Colon epithelia progenitors cDNA"
/note="Vector: pAMP1; 1st strand of cDNA was synthesized  
with reverse transcriptase and oligo(dT) beads, then cDNA  
was amplified by PCR using modified SMART primers. The  
final cDNA was cloned in pAMP1 vector in annealing  
reaction with uracil DNA Glycosylase (UDG). Library  
constructed by Y.Korshunova and M. Lovett. Library  
materials provided by J. Gordon lab."
```

ORIGIN

Query Match 1.3%; Score 50; DB 14; Length 269;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAAAA 3762
|||||


```

ORIGIN
treatment with UDG the cDNA was cloned in pAMP1 vector by
annealing."

Query Match      1.3%; Score 50; DB 14; Length 294;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
      |||||||
Db 225 TAAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 274

RESULT 57
CF384267      301 bp mRNA linear EST 27-AUG-2003
DEFINITION lac04h02.y1 SlEP Mus musculus cDNA 5', mRNA sequence.
ACCESSION CF384267
VERSION CF384267.1 GI:34327691
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
Tidwell,R., Clifton,S., Marra,M., Hillier,L., Pape,D., Martin,J.,
Wylie,T., Theising,B., Bowers,Y., Gibbons,M., Ritter,E., Bennet,J.,
Ronko,I., Tsagareishvili,R., Belaygorod,L., Grow,A., Maguire,L.,
Waterston,R. and Wilson,R.
Waterston,R. and Wilson,R.
WashU Stem cell EST Project
Unpublished (2002)
Contact: Jeff Gordon and Mike Lovett
WashU, Human Genetics Division
Washington University School of Medicine
Library materials provided by: Stappenbeck Th.S. and Gordon J.I.
Library constructed by: Korshunova Y. and Lovett M. DNA sequencing
by: Washington University Genome Sequencing Center For information
on obtaining a clone please contact: Rose Tidwell
(rtidwell@genetics.wustl.edu)
Putative full length read
vector to vector length is
Seq primer: -40RP from Gibco.
Location/Qualifiers
1.301
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/tissue_type="small intestinal epithelial progenitors"
/lab_host="DH5alpha Ultra Max cells (Invitrogen)"
/clone_lib="SlEP"
/note="Vector: pAMP1; The library was synthesized with
modified SMART primers with dUTP at the end. After
treatment with UDG the cDNA was cloned in pAMP1 vector by
annealing."

ORIGIN
Query Match      1.3%; Score 50; DB 14; Length 301;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
      |||||||
Db 228 TAAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 277

RESULT 58
AL045620      305 bp mRNA linear EST 04-SEP-2003
LOCUS AL045620
DEFINITION DKF2p4340045.r1 434 (synonym: htes3) Homo sapiens cDNA clone
ACCESSION DKF2p4340045
VERSION AL045620
KEYWORDS EST.

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Duesterhoeft,A., Lauber,J., Mewes,H.W., Gassenhuber,J. and
Wiemann,S.
EST (Duesterhoeft, et al.)
Unpublished (1999)
On Jul 9, 1999 this sequence version replaced gi:5433743.
Contact: MIPS
MIPS
Ingolstaedter Landstr 1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing
consortium of the German Genome Project.
No sl sequence available.
This clone (DKF2p4340045) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
Location/Qualifiers
1.305
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKF2p4340045"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="434 (synonym: htes3)"
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"

ORIGIN
Query Match      1.3%; Score 50; DB 9; Length 305;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
      |||||||
Db 184 TAAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 233

RESULT 59
CF754869      314 bp mRNA linear EST 16-OCT-2003
LOCUS lae07d02.y2 Colon epithelia progenitors cDNA Mus musculus cDNA 5',
DEFINITION mRNA sequence.
ACCESSION CF754869.1 GI:37698672
VERSION CF754869
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 314)
Tidwell,R., Clifton,S., Marra,M., Hillier,L., Pape,D., Martin,J.,
Wylie,T., Theising,B., Bowers,Y., Gibbons,M., Ritter,E., Bennet,J.,
Ronko,I., Tsagareishvili,R., Belaygorod,L., Grow,A., Maguire,L.,
Waterston,R. and Wilson,R.
WashU Stem cell EST Project
Unpublished (2002)
Contact: Jeff Gordon and Mike Lovett
WashU, Human Genetics Division
Washington University School of Medicine
1st strand of cDNA was synthesized with reverse transcriptase and
oligo(dT) beads, then cDNA was amplified by PCR using modified
SMART primers. The final cDNA was cloned in pAMP1 vector in
annealing reaction with Uracil DNA Glycosylase (UDG). Library
constructed by Y. Korshunova and M. Lovett. Library materials
provided by J. Gordon lab.
Putative full length read

```

vector to vector length is
 Seq primer: -40RP from Gibco.
 Location/Qualifiers
 source
 1..314
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /mol_type="mRNA"
 /tissue_type="colon epithelia progenitor cells"
 /lab_host="DH5alpha"
 /clone_lib="Colon epithelia progenitors cDNA"
 /note="Vector: pAMP1; 1st strand of cDNA was synthesized with reverse transcriptase and oligo(dT) beads, then cDNA was amplified by PCR using modified SMART primers. The final cDNA was cloned in pAMP1 vector in annealing reaction with Uracil DNA Glycosylase (UDG). Library constructed by Y. Korshunova and M. Lovett. Library materials provided by J. Gordon lab."

ORIGIN

Query Match 1.3%; Score 50; DB 14; Length 314;
 Best Local Similarity 100.0%; Pred. No. 1.6;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAAAA 321 bp mRNA linear EST 02-JUN-2003
 ||||| 3762
 Db 219 TAAATGAAAAA 268

RESULT 60
 CD422638 321 bp mRNA linear EST 02-JUN-2003
 LOCUS
 DEFINITION laa83g11.y1 SlEP Mus musculus cDNA 5', mRNA sequence.
 ACCESSION CD422638
 VERSION CD422638.1 GI:31328901
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)

ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 1 (bases 1 to 321)
 AUTHORS Tidwell, R., Clifton, S., Marra, M., Hillier, L., Pape, D., Martin, J., Wylie, I., Theising, B., Bowers, Y., Gibbons, M., Ritter, E., Bennet, D., Ronko, I., Teagareishvili, R., Belaygorod, L., Grow, A., Maguire, L., Waterston, R. and Wilson, R.
 TITLE WashU Stem cell EST Project
 JOURNAL Unpublished (2002)
 COMMENT Contact: Rose Tidwell
 WashU Stem cell EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 Library materials provided by: Stappenbeck Th.S. and Gordon J.I.
 Library constructed by: Korshunova Y. and Lovett M. DNA sequencing by: Washington University Genome Sequencing Center For information on obtaining a clone please contact: Rose Tidwell (rutidwell@genetics.wustl.edu)
 Putative full length read
 vector to vector length is
 Seq primer: -40RP from Gibco
 High quality sequence stop: 321.

FEATURES

source
 1..321
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /tissue_type="small intestinal epithelial progenitors"
 /lab_host="DH5alpha Ultra Max cells (Invitrogen)"
 /clone_lib="SlEP"
 /note="Vector: pAMP1; The library was synthesized with modified SMART primers with dUTP at the end. After treatment with UDG the cDNA was cloned in pAMP1 vector by

annealing."

ORIGIN
 Query Match 1.3%; Score 50; DB 14; Length 321;
 Best Local Similarity 100.0%; Pred. No. 1.6;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAAAA 329 bp mRNA linear EST 28-APR-1999
 ||||| 3762
 Db 219 TAAATGAAAAA 268

RESULT 61

AU034841
 LOCUS
 DEFINITION AU034841 Dictyostelium discoideum SL (H.Urushihara) Dictyostelium discoideum cDNA clone SLE570, mRNA sequence.

ACCESSION

AU034841 GI:3800265

KEYWORDS

Dictyostelium discoideum

SOURCE

Dictyostelium discoideum

ORGANISM

Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.

REFERENCE

1 (bases 1 to 329)

AUTHORS

Morio, T., Urushihara, H., Saito, T., Ugawa, Y., Mizuno, H., Yoshida, M., Yoshino, R., Mitra, B. N., Pi, M., Sato, T., Takemoto, K., Yasukawa, H., Williams, J., Maeda, M., Takeuchi, I., Ochiai, H. and Tanaka, Y.

TITLE

The Dictyostelium developmental cDNA project: generation and analysis of expressed sequence tags from the first-finger stage of development

development

DNA Res. 5 (6), 335-340 (1998)

99156227

10048482

Contact: Hideko Urushihara

Institute of Biological Sciences

University of Tsukuba

1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8572, Japan

Tel: 81-298-53-4664

Fax: 81-298-53-6614

Email: hideko@biol.tsukuba.ac.jp

PROJECT = 'Dictyostelium discoideum cDNA project in Japan'

POLYA=No.

Location/Qualifiers

1..329

/organism="Dictyostelium discoideum"

/mol_type="mRNA"

/strain="AX4"

/db_xref="taxon:44689"

/clone="SLE570"

/dev_stage="slug"

/clone_lib="Dictyostelium discoideum SL (H.Urushihara)"

ORIGIN

Query Match 1.3%; Score 50; DB 9; Length 329;

Best Local Similarity 100.0%; Pred. No. 1.5;

Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAAAA 3762

Db 273 TAAATGAAAAA 322

RESULT 62

BU767819

LOCUS

DEFINITION BU767819 SJE Schistosoma japonicum cDNA, mRNA sequence.

ACCESSION BU767819

VERSION BU767819.1 GI:28335135

KEYWORDS EST.

SOURCE Schistosoma japonicum

ORGANISM Schistosoma japonicum

Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;

Strigeidida; Schistosomatidae; Schistosoma.

REFERENCE 1 (bases 1 to 336)
 AUTHORS Hu, W., Yan, Q., Shen, D. K., Liu, F., Zhu, Z. D., Song, H. D., Xu, X. R., Wang, Z. J., Rong, Y. P., Zeng, L. C., Wu, J., Zhang, X., Wang, J. J., Xu, X. N., Wang, S. Y., Fu, G., Zhang, X. L., Wang, Z. Q., Brindley, P. J., McKusick, D. P., Xue, C. L., Peng, Z., Chen, Z., and Han, Z. G.
 TITLE Evolutionary and biomedical implications of a Schistosoma japonicum complementary DNA resource
 JOURNAL Nat. Genet. 35 (2), 139-147 (2003)
 MEDLINE 22879925
 COMMENT Contact: Zeguang Han
 Chinese National Human Genome Center at Shanghai
 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China
 Tel: 86-21-50801919 (ex. 45)
 Fax: 86-21-50801922
 Email: hanzg@chgc.sh.cn.
 FEATURES source
 1. .336
 Location/Qualifiers
 /organism="Schistosoma japonicum"
 /mol_type="mRNA"
 /db_xref="taxon:6182"
 /tissue_type="Whole egg"
 /dev_stage="egg"
 /lab_host="rabbits"
 /clone_lib="SJE"
 ORIGIN
 Query Match 1.3%; Score 50; DB 13; Length 336;
 Best Local Similarity 100.0%; Pred. No. 1.5;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3713 TAAATGAAA 3762
 DB 270 TAAATGAAA 319
 RESULT 63
 CF382593
 LOCUS lac88g07.y1 SLEP Mus musculus CDNA 5', mRNA sequence. EST 27-AUG-2003
 DEFINITION
 ACCESSION CF382593
 VERSION
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 346)
 Tidwell, R., Clifton, S., Marra, M., Hillier, L., Pape, D., Martin, J., Wylie, T., Theising, B., Bowers, Y., Gibbons, M., Ritter, E., Bennett, J., Ronko, I., Tsagarisshvili, R., Belaygorod, L., Grow, A., Maguire, L., Waterston, R. and Wilson, R.
 TITLE WashU Stem cell EST Project
 JOURNAL Unpublished (2002)
 COMMENT Contact: Jeff Gordon and Mike Lovett
 WashU, Human Genetics Division
 Washington University School of Medicine
 Library materials provided by: Stappenbeck Th.S. and Gordon J.I.
 Library constructed by: Korshunova Y. and Lovett M. DNA sequencing by: Washington University Genome Sequencing Center For information on obtaining a clone please contact: Rose Tidwell (rtidwellgenetics.wustl.edu)
 Putative full length read
 vector to vector length is
 Seq primer: -40RP from Gibco.
 Location/Qualifiers
 1. .346
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /tissue_type="small intestinal epithelial progenitors"
 /lab_host="DH5alpha Ultra Max cells (Invitrogen)"
 /clone_lib="SIEP"

/note="Vector: pAMP1; The library was synthesized with modified SMART primers with dUTP at the end. After treatment with UDG the cDNA was cloned in pAMP1 vector by annealing."
 ORIGIN
 Query Match 1.3%; Score 50; DB 14; Length 346;
 Best Local Similarity 100.0%; Pred. No. 1.5;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3713 TAAATGAAA 3762
 DB 225 TAAATGAAA 274
 RESULT 64
 BQ450975
 LOCUS PfEST0ab02c10.y1 Plasmodium falciparum 3D7 asexual cDNA Plasmodium falciparum 3D7 CDNA 5', mRNA sequence.
 DEFINITION
 ACCESSION BQ450975
 VERSION
 KEYWORDS EST.
 SOURCE Plasmodium falciparum 3D7
 ORGANISM Plasmodium falciparum 3D7
 Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 1 (bases 1 to 350)
 Tang, K., Cole, R., Chakrabarti, D., Haywood, R., Clifton, S., Pape, D., Marra, M., Hillier, L., Martin, J., Wylie, T., Dante, M., Theising, B., Bowers, Y., Gibbons, M., Ritter, E., Bennett, J., Jentes, E., Ronko, I., Tsagarisshvili, R., Belaygorod, L., Franklin, C., Carr, L., Grow, A., Maguire, L., Richey, J., Wadkins, J., Kennedy, S., Levinso, D., Waterston, R., Wilson, R. and Sibley, D.
 TITLE WashU Plasmodium EST Project
 JOURNAL Unpublished (2001)
 COMMENT Contact: L. David Sibley
 WashU Plasmodium EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@wustl.edu
 Library was constructed by Debopam Chakrabarti DNA sequencing by: Washington University Genome Sequencing Center For information on obtaining a clone please contact: L. David Sibley (sibley@bcm.wustl.edu), Washington University
 Seq primer: -40UP from Gibco.
 Location/Qualifiers
 1. .350
 /organism="Plasmodium falciparum 3D7"
 /mol_type="mRNA"
 /db_xref="taxon:36329"
 /lab_host="DH10B (GeneHog, Invitrogen, Inc.)"
 /clone_lib="Plasmodium falciparum 3D7 asexual cDNA"
 /note="Vector: pBluescript SK plus; Site 1: EcoRI; Site 2: XhoI; Library was constructed by Debopam Chakrabarti.
 Total RNA samples were isolated from mixed stage sporozoites (0.1%)-lysed P. falciparum 3D7 infected erythrocytes by the acidic guanidinium-phenol chloroform method. The poly A+ RNA was isolated by the polyAT-Tract mRNA isolation system (Promega, WI) using streptavidin Magnosphere particles. Directional cDNA libraries were constructed by oligo d(T) priming of poly(A)+ RNA (5mg) into EcoRI and XhoI sites of 1 ZapII vector using the Zap cDNA synthesis kit (Stratagene, CA). The average size of the cDNA inserts in the library was between 1.0 and 1.5kb. Clones were mass excised using the ExAssist helper phage (Stratagene), the phagemids were precipitated with PEG 8000 and extracted with phenol/chloroform. Phagemid DNA was electroporated into DH10B cells."

```

Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 296 TAAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 345

RESULT 65
BE048163 354 bp mRNA linear EST 08-JUN-2000
LOCUS T246905.Y1 NCI_CGAP_Brn52 Homo sapiens cDNA clone IMAGE:2291672 5',
DEFINITION mRNA sequence.
ACCESSION BE048163
VERSION BE048163.1 GI:8365141
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 354)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
R. Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LINL, send email to:
info@image.lnl.gov
Seq primer: -40RP from Gibco
High quality sequence stop: 297.
FEATURES
source
1..354
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2291672"
/tissue_type="tumor, 5 pooled (see description)"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Brn52"
/note="Organ: Brain; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; This library represents the normalized
version of NCI_CGAP_Brn35. Cloned unidirectionally.
Primer: Oligo dT. Average insert size 1.19 kb. Tumor
types include: meningioma, oligodendroglioma, astrocytoma
(grade II), medulloblastoma, astrocytoma (grade IV).
Constructed by Life Technologies."
ORIGIN
Query Match 1.3%; Score 50; DB 10; Length 354;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 215 TAAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 264

RESULT 66
BW306217 366 bp mRNA linear EST 11-NOV-2002
LOCUS BW306217 Nori Satoh unpublished cDNA library, heart Ciona
DEFINITION intestinalis cDNA clone cinfo12005 5', mRNA sequence.
ACCESSION BW306217
VERSION BW306217.1 GI:24886828
KEYWORDS EST.
SOURCE Ciona intestinalis

```

```

Ciona intestinalis
Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
Phlebobranchia; Cionidae; Ciona.
REFERENCE 1 (bases 1 to 366)
AUTHORS Satou, Y., Shin-I, T., Kohara, Y. and Satoh, N.
TITLE Expressed genes in Ciona intestinalis (2002c)
JOURNAL Unpublished (2002)
COMMENT Contact: Nori Satoh
Department of Zoology
Kyoto University
Sakyo-Ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: satoh@ascidian.zool.kyoto-u.ac.jp.
FEATURES
source
1..366
/organism="Ciona intestinalis"
/mol_type="mRNA"
/db_xref="taxon:7719"
/clone="cinfo12005"
/tissue_type="heart"
/clone_lib="Nori Satoh unpublished cDNA library, heart"
ORIGIN
Query Match 1.3%; Score 50; DB 13; Length 366;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 305 TAAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 354

RESULT 67
BG955855 387 bp mRNA linear EST 05-JUN-2001
LOCUS BG955855
DEFINITION 359509 MARC 1PIG Sus scrofa cDNA 5', mRNA sequence.
ACCESSION BG955855
VERSION BG955855.1 GI:14306096
KEYWORDS EST.
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (bases 1 to 387)
AUTHORS Fahrénkrug, S.C., Smith, T.P.L., Frenking, B.A., Cho, J., White, J.,
Vallet, J., Wise, T., Rohrer, G.A., Pertea, G., Sultana, R.,
Quackenbush, J. and Keefe, J.W.
TITLE Porcine gene discovery by normalized cDNA-library sequencing and
EST cluster assembly
JOURNAL Mamm. Genome 13 (8), 475-478 (2002)
MEDLINE 22213789
PUBMED 12226715
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.
PCR Primers
FORWARD: AGGAAACAGCTATGACCAT
BACKWARD: GTTTTCCAGTCACGCG
Plate: 124 row: 1 column: 6
Seq primer: ATTTAGGTGACACTATAG.
FEATURES
source
1..387
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/tissue_type="pooled"

```


using the SMART template was also desired. 15 PCR cycles were done using 1st strand and primers specific to SL1 leader sequence (SL1-NotI) and 3' end (XhoI-NO-dT). dscDNA was digested using XhoI/NotI, fractionated on Chroma-spin 400 columns (Clontech) and ligated to digested pGEM-11zf(+) plasmid. Chemically competent cells were used as host. This library was provided by David Bird's lab (david.bird@ncsu.edu), of North Carolina State University. Seq primer: Sp6

High quality sequence stop: 441.

FEATURES

source Location/Qualifiers

1..441
/organism="Meloiodogyne hapla"
/mol_type="mRNA"
/db_xref="taxon:6305"
/sex="female"
/tissue_type="whole organism"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="Meloiodogyne hapla female SL1 pGEM"
Note="Vector: pGEM-11zf(+) (Promega); Site 1: XhoI;
Site 2: NotI; Cloned unidirectionally. Poly (A)+ RNA was concentrated and purified using Dynabeads (Dyna) and mRNA eluted for 1st strand synthesis. 1st strand cDNA created using MMV RT(PowerScript, Clontech) and primed with oligo(dT) with XhoI site (primer CDSIII/3-XhoI) and 5' SMART anchor added using chimeric DNA-RNA oligo (SMART-NotI-r-GGG). SMART-NotI-r-GGG was used in case a PCR using the SMART template was also desired. 15 PCR cycles were done using 1st strand and primers specific to SL1 leader sequence (SL1-NotI) and 3' end (XhoI-NO-dT). dscDNA was digested using XhoI/NotI and ligated to digested Chroma-spin 400 columns (Clontech) and ligated to digested pGEM-11zf(+) plasmid. Chemically competent cells were used as host. This library was provided by David Bird's lab (david.bird@ncsu.edu), of North Carolina State University."

ORIGIN

Query Match 1.3%; Score 50; DB 14; Length 441;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAA 3762

Db 370 TAAATGAAA 419

RESULT 73

BM958905
LOCUS 443 bp mRNA linear EST 18-MAR-2002
DEFINITION PLATE 11 C09 05 AB1 GS Lambda-Triplex, 10 day germinating spore
Library Glomus intraradices cDNA, mRNA sequence.

ACCESSION BM958905

VERSION BM958905.1 GI:19542335

KEYWORDS EST.

SOURCE Glomus intraradices

ORGANISM Glomus intraradices

REFERENCE Eukaryota; Fungi; Glomeromycota; Glomeromycetes; Glomerales;

Glomeraceae; Glomus.

1 (bases 1 to 443)

Authors Bago, B., Zipfel, W., Williams, R.M., Jun, J., Arreola, R.,

Lammers, P.J., Pfeffer, P.E. and Shachar-Hill, Y.

Title Translocation and utilization of fungal storage lipid in the

arbuscular mycorrhizal symbiosis

Journal Plant Physiol. 128 (1), 108-124 (2002)

MEDLINE 21648947

PubMed 11788757

Comment Contact: Peter Lammers

New Mexico State University

Box 30001, MSC 3MLS Horseshoe Drive Las Cruces, NM 88003, USA

Tel: 505-646-3918

Fax: 505-646-6846

Email: plammers@nmsu.edu.

FEATURES

source

Location/Qualifiers
1..443
/organism="Glomus intraradices"
/mol_type="mRNA"
/db_xref="taxon:4876"
/tissue_type="germinating spore tissue"
/dev_stage="10 days after germination"
/lab_host="E. coli BM 25.8"
/clone_lib="GS Lambda-Triplex, 10 day germinating spore library"
Note="Vector: Lambda Triplex2; Site 1: SfiI; Site 2: SfiIIB; Spores were germinated in the presence of 1% CO₂. At 10 days, tissue was ground in liquid nitrogen with sand and mRNA isolated with a modified hot phenol/SDS method. The SMART-PCR method (Clontech) was used for cDNA synthesis from 1 ug of total RNA followed by directional cloning in SfiI digested Triplex2 vector. Plasmid subclones in pTriplex were recovered by cre-lox excision in E. coli strain BM25.8 and sequenced from the 5' end with the 5'Triplex sequencing primer (Tpx) and the 3' end with the T7 primer (T7). (sfiI= bp in Comment line of EST file is estimated insert size)"

ORIGIN

Query Match 1.3%; Score 50; DB 12; Length 443;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAA 3762

Db 361 TAAATGAAA 410

RESULT 74

C91260 457 bp mRNA linear EST 20-APR-1998
LOCUS C91260 Dictyostelium discoideum SS (H.Urushihara) Dictyostelium
DEFINITION discoideum cDNA clone SSU516, mRNA sequence.

ACCESSION C91260

VERSION C91260.1 GI:3060626

KEYWORDS EST.

SOURCE Dictyostelium discoideum

ORGANISM Dictyostelium discoideum

REFERENCE Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.

AUTHORS 1 (bases 1 to 457)

Morio, T., Urushihara, H., Saito, T., Ugawa, Y., Mizuno, H., Yoshida, M.,

Yoshino, R., Mitra, B.N., Pi, M., Sato, T., Takemoto, K., Yasukawa, H.,

Williams, J., Maeda, M., Takeuchi, I., Ochiai, H. and Tanaka, Y.

Developmental cDNA in Dictyostelium discoideum

Unpublished (1998)

Contact: Hideko Urushihara

Institute of Biological Sciences

University of Tsukuba

1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8572, Japan

Tel: 81-298-53-4664

Fax: 81-298-53-6614

Email: hideko@biol.tsukuba.ac.jp.

Location/Qualifiers

1..457

/organism="Dictyostelium discoideum"

/mol_type="mRNA"

/strain="AX4"

/db_xref="taxon:44689"

/clone="SSJ516"

/dev_stage="slug"

/clone_lib="Dictyostelium discoideum SS (H.Urushihara)"

ORIGIN

Query Match 1.3%; Score 50; DB 13; Length 457;

Best Local Similarity 100.0%; Pred. No. 1.1;

Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3712 ATAAATGAAA 3761

/clone="IMAGE:6846599"
 /tissue type="whole brain"
 /dev stage="1, 5, and 15 days newborn"
 /lab host="DH10B (T1 phage resistant)"
 /clone lib="NTH_BMAP_GKO"
 /note="Organ: Brain; Vector: pYX- Asc; Site_1: EcoR I; Site_2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is CGAATGGAAT. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Query Match 1.3%; Score 50; DB 14; Length 477;
 Best Local Similarity 100.0%; Pred. No. 1.1;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAAAA 483 bp mRNA linear EST 26-AUG-2003
 |||||
 Db 410 TAAATGAAAAA 459

RESULT 78

LOCUS B1376676 483 bp mRNA linear EST 26-AUG-2003
 DEFINITION BFLG3_000479 Amphioxus 5-6 hrs cDNA library (Name convention: BFLG or MPMPGp498) Branchiostoma floridae cDNA clone MPMPGp498B0638 5', mRNA sequence.

ACCESSION B1376676
 VERSION B1376676.1 GI:30911635

KEYWORDS EST.
 SOURCE Branchiostoma floridae (Florida lancelet)
 ORGANISM Branchiostoma floridae
 Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae; Branchiostoma.

REFERENCE 1 (bases 1 to 483)
 Panopoulou, G., Hennig, S., Groth, D., Krause, A., Poustka, A. J., Herwig, R., Vingron, M. and Lehrach, H.
 New evidence for genome-wide duplications at the origin of vertebrates using an amphioxus gene set and completed animal genomes

JOURNAL Genome Res. 13 (6A), 1056-1066 (2003)
 MEDLINE 22683279
 PUBMED 12799346

COMMENT Contact: Panopoulou G
 Laboratory 145, dept. Lehrach
 Max-Planck-Institut fuer Molekulare Genetik
 Ihnestr.63-73, D-14195 Berlin, Germany
 Tel: +49 30 8413 1235
 Fax: +49 30 8413 1128
 Email: panopoulou@molgen.mpg.de

The library was characterised by oligonucleotide fingerprinting (ONFP) to reduce sequencing redundancy. According to the ONFP procedure. Clones giving the same hybridisation pattern with a battery of 200 emer oligonucleotides are grouped into clusters. One clone per cluster is selected for sequencing. The size of each cluster is an indicator of the frequency of a transcript in the analysed library. The cluster size as well the coordinates of the rest of the clones assigned to the same fingerprint cluster as the clone from which the above EST is generated is available at the amphioxus project site at: <http://www.molgen.mpg.de/amphioxus/>
 Clones and filters are distributed via the Resource Center/Primary Database of the German Genome Project (<http://www.rzpd.de>)

PCR Primers

FORWARD: 5' CCCCAGGCTTACACTTATGCTTCGGCTCG 3' (ML13RSP)
 BACKWARD: 5' GCTATTACCCAGCTGGCGAAGGGGATG 3' (ML13SP)
 Insert Length: 1200 Std Error: 0.00
 Seq primer: 5' CCGGTCGGAATTCGGGT-3' pSport3/86
 High quality sequence stop: 483.

FEATURES

source Location/Qualifiers
 1..483
 /organism="Branchiostoma floridae"
 /mol type="mRNA"
 /db xref="taxon:7739"
 /clone="MPMPGp498B0638"
 /tissue type="whole embryo"
 /dev stage="5-6 hrs (gastrula stage)"
 /lab host="E.coli, XL1 blue"
 /clone lib="Amphioxus 5-6 hrs cDNA library (Name convention: BFLG or MPMPGp498)"
 /note="Vector: pSport1; Site 1: SalI, KpnI, EcoRI (5'); Site 2: NotI, BamHI, HindIII (3'); OligodT primed and directionally cloned in pSport1 vector using a NotI (5'-pGACTAGTTCTAGTCGCGAGCGCGCCC (T)15-3' and a SalI 5'-TCGACCCAGCGGTCCG-3' adapters (Gibco BRL)."

ORIGIN

Query Match 1.3%; Score 50; DB 12; Length 483;
 Best Local Similarity 100.0%; Pred. No. 1;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAAAA 503 bp mRNA linear EST 24-JUN-2003
 |||||
 Db 358 TAAATGAAAAA 3762

RESULT 79

LOCUS CD681869
 DEFINITION

ACCESSION CD681869
 VERSION CD681869.1 GI:32183635
 KEYWORDS Meloidogyne chitwoodi
 SOURCE Meloidogyne chitwoodi
 ORGANISM Meloidogyne chitwoodi

REFERENCE 1 (bases 1 to 503)
 McCarter, J., Clifton, S., Chiapelli, B., Pape, D., Martin, J., Wyllie, T., Dante, M., Marra, M., Hillier, L., Kucaba, T., Theising, B., Bowers, Y., Gibbons, M., Ritter, E., Bennett, J., Franklin, C., Tsagaris, V., Ronko, I., Kennedy, S., Maguire, L., Beck, C., Underwood, K., Steptoe, M., Allen, M., Person, B., Swaller, T., Harvey, N., Schurk, R., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.

TITLE The Washington Univ. Nematode EST Project, 1999
 JOURNAL Unpublished (1999)
 COMMENT Contact: McCarter JP
 The Washington Univ. Nematode EST Project, 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu

The library was constructed by Claire Murphy and Dr. Makedonka Dautova at Washington University, St. Louis
 (mdautova@watson.wustl.edu). Oligo(dT)-SL1 PCR based library.
 Meloidogyne chitwoodi egg cDNA PCR products of size >400 nucleotides containing SL1 on the 5' end and oligo(dT) on the 3' end were non-directionally cloned into pCRII-TOPO (Invitrogen) following the TOPO TA cloning protocol. Eggs were provided by Dr. David Bird of North Carolina State University, Raleigh, NC (david.bird@ncsu.edu).
 Seq primer: SL1 primer
 High quality sequence stop: 503.

FEATURES
source

Location/Qualifiers
1. .503
/organism="Meloioogyne chitwoodi"
/mol_type="mRNA"
/db_xref="taxon:59747"
/dev_stage="eggs"
/lab_host="DH10B"
/clone_lib="Meloioogyne chitwoodi egg SL1 TOPO v1"
/note="Vector: PCR11-TOPO (Invitrogen); Site_1: EcoRI;
Site_2: EcoRI; The library was constructed by Claire
Murphy and Dr. Makedonka Bautova at Washington University,
St. Louis. Oligo(dT)-SL1 PCR based library. Meloioogyne
chitwoodi egg cDNA PCR products of size >400 nucleotides
containing SL1 on the 5' end and oligo(dT) on the 3' end
were non-directionally cloned into PCR11-TOPO(Invitrogen)
following the TOPO TA cloning protocol. Eggs were provided
by Dr. David Bird of North Carolina State University,
Raleigh, NC (david_bird@ncsu.edu)."

ORIGIN

Query Match 1.3%; Score 50; DB 14; Length 503;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3713 TAAATGAAAAA 3762
DB 415 TAAATGAAAAA 464

RESULT 80

BM938356
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PubMed
COMMENT

BM938356 515 bp mRNA linear EST 13-MAR-2002
UI-M-BZ1-bfw-o-15-0-UI-r1 NIH BMAP MH12 S1 Mus musculus cDNA clone
UI-M-BZ1-bfw-o-15-0-UI 5', mRNA sequence.
BM938356
BM938356.1 GI:19397508
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 515)
Ronald, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
9704477
889548
Contact: Chin, H
National Institute of Mental Health
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
20892-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890
Email: mEST@mail.nih.gov
cDNA library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
The following repetitive elements were found in this cDNA
sequence: 83-187, >(TAA)nSimple_repeat 105-189,
>POLY(A)Simple_repeat
Seq primer: M13 REVERSE.

FEATURES
source

Location/Qualifiers
1. .515
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UI-M-BZ1-bfw-o-15-0-UI"
/dev_stage="27-32 days"
/lab_host="DH10B (Life Technologies)"

/clone_lib="NIH BMAP MH12 S1"
/note="Vector: pTT3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; The
NIH BMAP MH12 S1 library is a subtracted library derived
from NIH BMAP MH12. NIH BMAP MH12 is a library derived
from mouse hippocampus tissue. For a detailed description
of the library from which this clone was derived, please
visit our web site at brainest.eng.uowa.edu."

ORIGIN

Query Match 1.3%; Score 50; DB 12; Length 515;
Best Local Similarity 100.0%; Pred. No. 0.98;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3713 TAAATGAAAAA 3762
DB 104 TAAATGAAAAA 153

RESULT 81

AM464287
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AM464287 562 bp mRNA linear EST 24-FEB-2000
BP230015A20F2 Soares normalized bovine placenta Bos taurus cDNA
clone BP230015A20F2 5', mRNA sequence.
AM464287
AM464287.1 GI:7034455
EST
Bos taurus (cow)
Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 562)
Lewin, H.A., Soares, M.B., Rebeiz, M., Pardinas, J., Liu, L. and
Larson, J.H.
Bovine ESTs
Unpublished (2000)
Contact: Lewin, H. A.
W. M. Keck Center for Comparative and Functional Genomics
University of Illinois at Urbana-Champaign
340 Edward R. Madigan Laboratory, 1201 W. Gregory Dr., Urbana, IL
61801, USA
Tel: 217 333 5998
Fax: 217 244 5617
Email: h-lewin@uiuc.edu

Funding for cattle EST sequencing was provided by the USDA National
Research Initiative, Animal Genome Resource Grant AG 99-3205-8534
to H. A. Lewin and J. E. Womack. Base Calling/Quality Scores: PHRED
from Washington University Genome Center. Vector trimmi g:
Crossmatch from Washington University Genome Center PHRAP suite.
Sequences submitted are vector free and at least 200 bp in length.
PCR Primers
FORWARD: TAAACGACTCACTATAGGG
BACKWARD: ATTAACCTCTACTAAAG

Insert Length: 562 Std Error: 0.00
Plate: BP230015A20 Row: F Column: 2
Seq primer: ACGGATACAAATTCACACAGGA
High quality sequence stop: 562.
Location/Qualifiers
1. .562
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/clone="BP230015A20F2"
/sex="female"
/lab_host="DH10B"
/clone_lib="Soares normalized bovine placenta"
/note="Organ: placenta; Vector: pTT3Pac; Site_1: EcoRI;
Site_2: NotI; The cDNA library was contributed by the
Soares laboratory and it was constructed and normalized
as described by Bonaldo, M.F., Lennon, G. and Soares,
M.B. (1996), Genome Research 6(9): 791-806."

FEATURES
source

Location/Qualifiers
1. .562
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/clone="BP230015A20F2"
/sex="female"
/lab_host="DH10B"
/clone_lib="Soares normalized bovine placenta"
/note="Organ: placenta; Vector: pTT3Pac; Site_1: EcoRI;
Site_2: NotI; The cDNA library was contributed by the
Soares laboratory and it was constructed and normalized
as described by Bonaldo, M.F., Lennon, G. and Soares,
M.B. (1996), Genome Research 6(9): 791-806."

ORIGIN


```

REFERENCE
AUTHORS
  Query Match      1.3%; Score 50; DB 10; Length 562;
  Best Local Similarity 100.0%; Pred. No. 0.9;
  Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
DB 501 TAAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 550

RESULT 82
LOCUS BX361013 602 bp mRNA linear EST 05-MAY-2003
DEFINITION BX361013 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA
ACCESSION clone CSODI078YXN19 5-PRIME, mRNA sequence.
VERSION BX361013
KEYWORDS BX361013.1 GI:30384484
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 602)
Li.W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 7394.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
Cgi-bin/cluster.cgi?seq=CSODI078CG100P1&cluster=7394.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CSODI078CG100P1.

FEATURES
Location/Qualifiers
1..602
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODI078YXN19"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match      1.3%; Score 50; DB 13; Length 602;
Best Local Similarity 100.0%; Pred. No. 0.84;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
DB 386 TAAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 435

RESULT 83
LOCUS CF979621 621 bp mRNA linear EST 25-NOV-2003
DEFINITION CF979621 y1 Meloidogyne hapla female SL1 pGEM Meloidogyne hapla CDNA
ACCESSION 5', mRNA sequence.
VERSION CF979621
KEYWORDS CF979621.1 GI:38512670
SOURCE EST.
ORGANISM Meloidogyne hapla
Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;
Tylenchoidea; Heteroderidae; Meloidogyninae; Meloidogyne.

```

```

REFERENCE
AUTHORS
  1 (bases 1 to 621)
  McCarter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J.,
  Wylie,T., Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B.,
  Bowers,Y., Gibbons,M., Ritter,E., Bennett,J., Franklin,C.,
  Tsagarisakivi,I., Ronko,I., Kennedy,S., Maguire,L., Beck,C.,
  Underwood,K., Steptoe,M., Allen,M., Person,B., Swaller,T.,
  Harvey,N., Schurk,R., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
  McCann,R., Waterston,R. and Wilson,R.
  The Washington Univ. Nematode EST Project, 1999
  Unpublished (1999)
  Contact: McCarter JP
  The Washington Univ. Nematode EST Project, 1999
  Washington University School of Medicine
  4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
  Tel: 314 286 1800
  Fax: 314 286 1810
  Email: est@wustl.wustl.edu
  Cloned unidirectionally. Poly (A) + RNA was concentrated and
  purified using Dynabeads (Dyna) and mRNA eluted for 1st strand
  synthesis. 1st strand cDNA created using MMV RT(PowerScript,
  Clontech) and primed with oligo(dT) with XhoI site (primer,
  CDSIII/3-XhoI) and 5' SMART anchor added using chimeric DNA-RNA
  oligo (SMART-NotI-r-GGG). SMART-NotI-r-GGG was used in case a PCR
  using the SMART template was also desired. 15 PCR cycles were done
  using 1st strand and primers specific to SL1 leader sequence
  (SL1-NotI) and 3' end (XhoI-No-dT). dscDNA was digested using
  XhoI/NotI, fractionated on Chroma-spin 400 columns (Clontech) and
  ligated to digested pGEM -11zf(+) plasmid. Chemically competent
  cells were used as host. This library was provided by David Bird's
  lab (david.bird@ncsu.edu), of North Carolina State University.
  Seq primer: Sp6.

FEATURES
Location/Qualifiers
1..621
/organism="Meloidogyne hapla"
/mol_type="mRNA"
/db_xref="taxon:6305"
/sex="female"
/tissue_type="whole organism"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="Meloidogyne hapla female SL1 pGEM"
/note="Vector: pGEM-11zf(+) (Promega); Site 1: XhoI;
Site 2: NotI; Cloned unidirectionally, Poly (A) + RNA was
concentrated and purified using Dynabeads (Dyna) and mRNA
eluted for 1st strand synthesis. 1st strand cDNA created
using MMV RT(PowerScript, Clontech) and primed with
oligo(dT) with XhoI site (primer CDSIII/3-XhoI) and 5'
SMART anchor added using chimeric DNA-RNA oligo
(SMART-NotI-r-GGG). SMART-NotI-r-GGG was used in case a
PCR using the SMART template was also desired. 15 PCR
cycles were done using 1st strand and primers specific to
SL1 leader sequence (SL1-NotI) and 3' end (XhoI-No-dT).
dscDNA was digested using XhoI/NotI, fractionated on
Chroma-spin 400 columns (Clontech) and ligated to digested
pGEM -11zf(+) plasmid. Chemically competent cells were
used as host. This library was provided by David Bird's
lab (david.bird@ncsu.edu), of North Carolina State
University."

ORIGIN
Query Match      1.3%; Score 50; DB 14; Length 621;
Best Local Similarity 100.0%; Pred. No. 0.82;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
DB 371 TAAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 420

RESULT 84
LOCUS BC049569 637 bp mRNA linear HTC 08-APR-2003
DEFINITION BC049569 MUS musculus, RIKEN CDNA 1700054O13 gene, clone IMAGE:6772969,

```


REFERENCE
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Len Zon, Harvard
 cDNA Library Preparation: Open Biosystems
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLM14822 row: d column: 06
 High quality sequence stop: 103.

FEATURES
 source
 1..780
 /organism="Danio rerio"
 /mol_type="mRNA"
 /db_xref="taxon:7955"
 /clone="IMAGE:7050776"
 /tissue_type="whole body"
 /lab_host="DH10B"
 /clone_lib="NIH_ZGC_10"
 /note="Vector: pExpress1; Site 1: NotI; Site 2: EcoRV;
 Bulk tissue was collected from a whole adult individual
 from the Tuebingen strain. 1st strand cDNA was primed with
 a Not I - oligo(dT) primer, double-stranded cDNA was
 cloned into the Not I and EcoRV sites of pExpress-1.
 Library was size-selected for >1 kb fragments. A
 normalized version of this library is also available
 (NIH_ZGC 7). Library was constructed by Open Biosystems
 (Huntsville, AL)."

ORIGIN
 Query Match 1.3%; Score 50; DB 14; Length 780;
 Best Local Similarity 100.0%; Pred. No. 0.65;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3713 TAAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
 Db 61 TAAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 12

RESULT 86
 BUS98654
 LOCUS
 DEFINITION
 AGENCOURT_8822419 NIH_MGC_142 Homo sapiens cDNA clone IMAGE:6457561
 5', mRNA sequence.
 ACCESSION
 BUS98654
 VERSION
 BUS98654.1 GI:23250413
 KEYWORDS
 EST.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: NCI
 cDNA Library Preparation: Michael Brownstein Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLCW2639 row: k column: 02

REFERENCE
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: NCI
 cDNA Library Preparation: Michael Brownstein Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLCW2639 row: k column: 02

REFERENCE
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Len Zon, Harvard
 cDNA Library Preparation: Open Biosystems
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLM14822 row: d column: 06
 High quality sequence stop: 103.

FEATURES
 source
 1..780
 /organism="Danio rerio"
 /mol_type="mRNA"
 /db_xref="taxon:7955"
 /clone="IMAGE:7050776"
 /tissue_type="whole body"
 /lab_host="DH10B"
 /clone_lib="NIH_ZGC_10"
 /note="Vector: pExpress1; Site 1: NotI; Site 2: EcoRV;
 Bulk tissue was collected from a whole adult individual
 from the Tuebingen strain. 1st strand cDNA was primed with
 a Not I - oligo(dT) primer, double-stranded cDNA was
 cloned into the Not I and EcoRV sites of pExpress-1.
 Library was size-selected for >1 kb fragments. A
 normalized version of this library is also available
 (NIH_ZGC 7). Library was constructed by Open Biosystems
 (Huntsville, AL)."

ORIGIN
 Query Match 1.3%; Score 50; DB 14; Length 780;
 Best Local Similarity 100.0%; Pred. No. 0.65;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3713 TAAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
 Db 61 TAAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 12

RESULT 86
 BUS98654
 LOCUS
 DEFINITION
 AGENCOURT_8822419 NIH_MGC_142 Homo sapiens cDNA clone IMAGE:6457561
 5', mRNA sequence.
 ACCESSION
 BUS98654
 VERSION
 BUS98654.1 GI:23250413
 KEYWORDS
 EST.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: NCI
 cDNA Library Preparation: Michael Brownstein Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLCW2639 row: k column: 02

High quality sequence stop: 403.

FEATURES

source

1. .844
Location/Qualifiers
/organism="Homo sapiens"
/db_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6457561"
/tissue_type="mixed (pool of 40 RNAs)"
/lab_host="DH10B (TI-phage-resistant)"
/clone_lib="NIH_MGC_142"
/note="Vector: pDNR-LIB; Site 1: SfiI (ggccattatggcc); Site 2: SfiI (ggccctctggcc); Double-stranded cDNA was prepared from a pool of 40 cell line polyA+ RNAs (bladder - 2%, blood - 33.4%, brain - 5.6%, breast - 12.5%, colon - 4%, connective tissue - 1.4%, eye - 1%, intestine - 2.6%, kidney - 2.2%, liver - 5.7%, lung - 10.8%, NK-cell - 5.2%, ovary - 4%, pharynx - 2.5%, prostate - 4.3%, salivary gland - 1.3%, and skin - 2.3%). 5' and 3' adaptors were used in cloning as follows:
5'-AGACGTGTTATCAACGACGATGCGCATACGCCGGG-3' and 5'-ATTCTAGGCGCGGCGGCGGACATG-DT(30)NN-3'. Full-length enriched library was constructed using the Clontech Creator SMART kit and size-selected to contain the >0.5 kb size fraction (other fractions present in NIH_MGC_141). Library created in the laboratory of M. Brownstein (NIH, NIH). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 1.3%; Score 50; DB 13; Length 844;
Best Local Similarity 100.0%; Pred. No. 0.6;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3713 TAAATGAAAAA 849 bp mRNA linear EST 22-NOV-2000
Db 363 TAAATGAAAAA 412

RESULT 87
BF339838
LOCUS
DEFINITION
602038845F1 NCI_CGAP_Brn64 Homo sapiens cDNA clone IMAGE:4186834
5', mRNA sequence.

ACCESSION
BF339838
VERSION
BF339838.1 GI:11286297
KEYWORDS
EST.

SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 849)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)

COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: David N. Louis, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

Plate: LLN9508 row: i column: 11
High quality sequence stop: 568.
Location/Qualifiers

FEATURES

source

1. .849
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4186834"
/tissue_type="gliblastoma with EGFR amplification"
/lab_host="DH10B (TI phage-resistant)"
/clone_lib="NCI_CGAP_Brn64"

ORIGIN

Query Match 1.3%; Score 50; DB 10; Length 849;
Best Local Similarity 100.0%; Pred. No. 0.6;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3713 TAAATGAAAAA 865 bp mRNA linear EST 29-SEP-2000
Db 485 TAAATGAAAAA 534

RESULT 88

BE913663

LOCUS

DEFINITION

601669014F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:3969146 5',

mRNA sequence.

ACCESSION

BE913663

VERSION

BE913663.1 GI:10411508

KEYWORDS

EST.

SOURCE

Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE

1 (bases 1 to 865)

AUTHORS

NIH-MGC http://mgc.nci.nih.gov/.

TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLN9146 row: k column: 03

High quality sequence stop: 406.

Location/Qualifiers

1. .865

/organism="Mus musculus"

/mol_type="mRNA"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone="IMAGE:3969146"

/tissue_type="tumor, biopsy sample"

/dev_stage="10 months, virgin"

/lab_host="DH10B"

/clone_lib="NCI_CGAP Mam1"

/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;

Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.

Library constructed by Life Technologies. Investigator

providing samples: Gilbert Smith, NIH"

/note="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.57 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

Query Match 1.3%; Score 50; DB 10; Length 865;
Best Local Similarity 100.0%; Pred. No. 0.59;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3713 TAAATGAAAAA 883 bp mRNA linear EST 05-MAY-2003
Db 353 TAAATGAAAAA 402

RESULT 89

EX346858

LOCUS

DEFINITION

EX346858 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA

clone CSODI005YF20 5-PRIME, mRNA sequence.

```

ACCESSION BX346858
VERSION BX346858.1 GI:30379137
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 883)
JOURNAL Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
COMMENT Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0AT0172512_T01626_1.
Location/Qualifiers
1. .883
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1005YF20"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match 1.3%; Score 50; DB 13; Length 883;
Best Local Similarity 100.0%; Pred. No. 0.58;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3713 TAAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 28 TAAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 77

RESULT 90
LOCUS AU051237 892 bp mRNA linear EST 12-JUL-2000
DEFINITION AU051237 Sugano mouse brain mmb Mus musculus cDNA clone MNCB-1712
5', mRNA sequence.
ACCESSION AU051237
VERSION AU051237.1 GI:4434246
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
REFERENCE Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS Sasaki, M., Suzuki, K., Watanabe, M., Imai, J., Shibui, A., Yoshida, K.,
Hata, H., Yamaguchi, R., Tateyama, S. and Sugano, S.
TITLE Construction of mouse full length-enriched cDNA libraries by
oligo-capping
JOURNAL Unpublished (1999)
COMMENT Contact: Katsuyuki Hashimoto
Division of Genetic Resources
National Institute of Infectious Diseases
2-1-1, Toyama 1-chome, Shinjuku-ku, Tokyo 162-8640, Japan
Email: khashi@nih.go.jp
URL: http://www.nih.go.jp/yoken/genbank/
Seq primer: 5' end primer: CTTCTGCTCTAAAGCTGG
POLYA=No.
Location/Qualifiers
1. .892
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
FEATURES
source
/db_xref="taxon:10090"
/clone="MNCB-1712"
/sex="female"
/dev_stage="adult"
/lab_host="TOP10"
/clone_lib="Sugano mouse brain mmb"
/notes="Organ: brain; Vector: pME18S-FL3; 1st strand cDNA
was primed with an oligo(dT) primer
ATGTGGCTTTTTTTTTTTTTT; double-stranded cDNA was
ligated to a DraIII adaptor (TGTTGGCTACTGG), digested and
cloned into distinct DraIII sites of the pME18S-FL3. XhoI
sites just outside the DraIII sites can be used to isolate
the cDNA insert. Size selection was performed to exclude
fragments <1.5 kb. Library was constructed by Sugano et
al. (University of Tokyo, Institute of Medical Science).
Custom primer used for sequencing: 5' end primer
[CTTCTGCTCTAAAGCTGG], 3' end primer
[CGACCTGCAGCTCGAGCACA]"
ORIGIN
Query Match 1.3%; Score 50; DB 9; Length 892;
Best Local Similarity 100.0%; Pred. No. 0.57;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3713 TAAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 116 TAAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 165

RESULT 91
LOCUS BX325997 916 bp mRNA linear EST 01-MAY-2003
DEFINITION BX325997 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1082YI06 5-PRIME, mRNA sequence.
ACCESSION BX325997
VERSION BX325997.1 GI:30309528
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 7394.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0AT0242C12_T02299_1&cluster=7394.f.
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0AT0242C12_T02299_1.
Location/Qualifiers
1. .916
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1082YI06"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match 1.3%; Score 50; DB 13; Length 916;
Best Local Similarity 100.0%; Pred. No. 0.56;

```

```
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3713 TAAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
    |||||
Db 44 TAAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 93

RESULT 92
BX369283
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BX369283 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA
clone CSODI054YML7 5-PRIME, mRNA sequence.

ACCESSION BX369283.1 GI:30441388
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BX369283 Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 940)
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4928.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0BAF027ZH10_AF02577.1&cluster=4928.r.
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ invitrogen Corporation 1500
Faraday Avenue Genoscope sequence ID : CS0BAF027ZH10_AF02577.1.
Location/Qualifiers
1..940
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODI054YML7"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/Note="5' end strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 1.3%; Score 50; DB 13; Length 940;
Best Local Similarity 100.0%; Pred.No. 0.54;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3713 TAAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
    |||||
Db 209 TAAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 258

RESULT 93
BU945029
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BU945029 Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 942)
AGENCOURT 10546264 NIH_MGC_141 Homo sapiens CDNA clone
IMAGE:6733277 5', mRNA sequence.

ACCESSION BU945029
VERSION BU945029.1 GI:24133848
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BU945029 Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 942)
National Institutes of Health, Mammalian Gene Collection (MGC)
```

```
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: NCI
CDNA Library Preparation: Michael Brownstein Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/ILNL at:
http://image.llnl.gov
Plate: LLCM3062 row: k column: 04
High quality sequence stop: 352.
Location/Qualifiers
1..942
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6733277"
/tissue_type="mixed (pool of 40 RNAs)"
/lab_host="DH10B (T1-phage-resistant)"
/clone_lib="NIH_MGC_141"
/Note="Vector: pDNR-LIB; Site 1: SfiI (ggcattatggcc);
Site 2: SfiI (ggcgctcgcc); Double-stranded cDNA was
prepared from a pool of 40 cell line polyA+ RNAs (bladder
- 2%, blood - 33.4%, brain - 5.6%, breast - 12.5%, colon -
4%, connective tissue - 1.4%, eye - 1%, intestine - 2.6%,
kidney - 2.2%, liver - 5.7%, lung - 10.8%, NK-cell - 5.2%,
ovary - 4%, pharynx - 2.5%, prostate - 4.3%, salivary
gland - 1.3%, and skin - 2.3%). 5' and 3' adaptors were
used in cloning as follows:
5'-AAGCAGTGTATCAACGAGTGGCATTACGCGGG-3' and
5'-ATTAGAGCGCGAGCGCGCCATG-DT(30)NN-3'. Full-length
enriched library was constructed using the Clontech
Creator SMART kit and size-selected to contain the 0.2-0.5
kb size fraction (other fractions present in NIH_MGC 142).
Library created in the laboratory of M. Brownstein (NIH,
NIH). Note: this is a NIH_MGC library."

ORIGIN
Query Match 1.3%; Score 50; DB 13; Length 942;
Best Local Similarity 100.0%; Pred.No. 0.54;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3713 TAAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
    |||||
Db 245 TAAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 294

RESULT 94
BX326372
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BX326372 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA
clone CSODI082YI06 5-PRIME, mRNA sequence.

ACCESSION BX326372
VERSION BX326372.1 GI:30344535
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BX326372 Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 942)
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 7394.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0BAA018ZE02_CS01646_1&cluster=7394.f.
```

Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0BAF0162E02_CS01646_1.
Location/Qualifiers
1. .942
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI0821I06"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 1.3%; Score 50; DB 13; Length 942;
Best Local Similarity 100.0%; Pred.No.0.54;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3713 TAAATGAAA 3762
|||
Db 122 TAAATGAAA 171

RESULT 95
BX369284

LOCUS BX369284 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS0DI054YM17 5-PRIME, mRNA sequence.
ACCESSION BX369284
VERSION BX369284.1 GI:30441390
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 952)
Li, W.B., Gruber, C., Jesses, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4928.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0BAF027ZH10_AF02577_2&cluster=4928.r.
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0BAF027ZH10_AF02577_2.
Location/Qualifiers
1. .952
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI054YM17"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 1.3%; Score 50; DB 13; Length 952;
Best Local Similarity 100.0%; Pred.No.0.53;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3713 TAAATGAAA 3762

Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM12220 row: g column: 16
 High quality sequence start: 28
 High quality sequence stop: 412.

FEATURES
 source
 1. 994
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5534247"
 /tissue_type="leiomyosarcoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 71"
 /note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 2.1 kb."

ORIGIN
 Query Match 1.3%; Score 50; DB 12; Length 994;
 Best Local Similarity 100.0%; Pred. No. 0.51;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAAAA 3762
 |||||
 Db 255 TAAATGAAAAA 304

RESULT 98
 BG254352
 LOCUS
 DEFINITION
 602368921F1 NTH_MGC_91 Homo sapiens cDNA clone IMAGE:4477133 5', mRNA sequence.
 ACCESSION
 BG254352
 VERSION
 BG254352.1 GI:12764168
 KEYWORDS
 EST.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1023)
 NIH-MGC <http://mgi.nci.nih.gov/>
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: DCTD/DTF
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Cloned unidirectionally; Oligo-dT primed.
 Average insert size 1.4 kb. Library enriched for full-length clones and constructed by Life Technologies.

FEATURES
 source
 1. 1023
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4477133"
 /tissue_type="adenocarcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 91"
 /note="Organ: Prostate; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; Oligo-dT primed. Average insert size 1.4 kb. Library enriched for full-length clones and constructed by Life Technologies.

Note: this is a NIH_MGC Library."

ORIGIN
 Query Match 1.3%; Score 50; DB 12; Length 1023;
 Best Local Similarity 100.0%; Pred. No. 0.5; 0; Indels 0; Gaps 0;
 Matches 50; Conservative 0; Mismatches 0;

QY 3713 TAAATGAAAAA 3762
 |||||
 Db 139 TAAATGAAAAA 188

RESULT 99
 BG105078
 LOCUS
 DEFINITION
 602312366F1 NTH_MGC_84 Homo sapiens cDNA clone IMAGE:4422088 5', mRNA sequence.
 ACCESSION
 BG105078
 VERSION
 BG105078.1 GI:12598924
 KEYWORDS
 EST.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1041)
 NIH-MGC <http://mgi.nci.nih.gov/>
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Cloned unidirectionally; Oligo-dT primed.
 Average insert size 1.229 kb. Library enriched for full-length clones and constructed by Life Technologies.

FEATURES
 source
 1. 1041
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4422088"
 /tissue_type="adrenal cortex carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 84"
 /note="Organ: adrenal gland; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; Oligo-dT primed. Average insert size 1.229 kb. Library enriched for full-length clones and constructed by Life Technologies.

ORIGIN
 Query Match 1.3%; Score 50; DB 12; Length 1041;
 Best Local Similarity 100.0%; Pred. No. 0.49; 0; Indels 0; Gaps 0;
 Matches 50; Conservative 0; Mismatches 0;

QY 3713 TAAATGAAAAA 3762
 |||||
 Db 351 TAAATGAAAAA 400

RESULT 100
 BU526227
 LOCUS
 DEFINITION
 AGENCOURT 10173174 NIH MGC 144 Mus musculus cDNA clone IMAGE:6534925 5', mRNA sequence.
 ACCESSION
 BU526227
 VERSION
 BU526227.1 GI:22836668
 KEYWORDS
 EST.
 SOURCE
 Mus musculus (house mouse)

ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 1070)
 AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Dr. Michael Brownstein
 CDNA Library Preparation: Michael Brownstein Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: L1CM2693 row: n column: 13
 High quality sequence stop: 422.

FEATURES
 Location/Qualifiers
 1..1070
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:6534925"
 /lab_host="DH10B (Ti-phage-resistant)"
 /clone_lib="NIH_MGC_144"
 /note="Organ: Brain; Vector: pDNR-LIB; Site 1: SfiI
 (ggccataaggcc); Site 2: SfiI (ggcgctcgcc); cDNA made
 by oligo-dt priming and directionally cloned. 5' and 3'
 adaptors were used in cloning as follows:
 5'-AAGCAGTGTATCAACGACGATGCGCATTCAGCGCGG-3' and
 5'-ATTCTAGAGCGGCGGCGGCGGCGG-3'. Full-length
 enriched library was constructed using the Clontech
 Creator SMART kit and size-selected to contain the 0.2-0.5
 kb size fraction (other fractions present in NIH_MGC_143).
 Library created in the laboratory of M. Brownstein (NIH).
 NIH). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 1.3%; Score 50; DB 13; Length 1070;
 Best Local Similarity 100.0%; Pred.No. 0.48;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAA 3762
 |||
 Db 525 TAAATGAAA 574

Search completed: April 23, 2004, 11:01:47
 Job time : 8962 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 22, 2004, 16:09:21 ; Search time 2165 Seconds
(without alignments)
2702.683 Million cell updates/sec

Title: US-10-005-907-2

Perfect score: 715

Sequence: 1 MGNYLRLKLSCLGENQKPK.....SVSRPCSTHEHDYEVVFP 135

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DSV=xlp
-Q=/cgn2 1/USPTO.spool_p/US10005907/runat_21042004_113754_11432/app.query.fasta_1.327
-DB=GenEmbl -QFMT=fastap -SUFFIX=sgs -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNIT=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAFSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10005907@cgn 1 1 5265 @runat_21042004_113754_11432 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:

1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_io:*
27: em_sts:*
28: em_un:*

29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rdi:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_ey:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	715	100.0	2648	9	BC024174 Homo sapi
2	715	100.0	3762	6	AX505122 Sequence
3	711	99.4	1709	9	AX123798 Homo sapi
4	634	88.7	1924	9	AB060908 Macaca fa
5	584	81.7	4032	9	AX126682 Homo sapi
6	581	81.3	4136	9	AX124520 Homo sapi
7	430	60.1	141268	9	AC074365 Homo sapi
8	430	60.1	185467	9	AL606804 Human DNA
9	169.5	23.7	1532	9	BC030506 Homo sapi
10	169.5	23.7	1654	9	AF521911 Homo sapi
11	169.5	23.7	3270	9	AX212446 Homo sapi
12	159	22.2	1643	6	AX740457 Sequence
13	128	17.9	1672	10	U13263 Mus musculu
14	126	17.6	8218	9	AX125521 Homo sapi
15	126	17.6	81001	9	AC128688 Homo sapi
16	126	17.6	203405	2	AC024964 Homo sapi
17	126	17.6	214324	2	AC110870 Homo sapi
18	112	15.7	86826	3	PFWAL3P5 Plasmodi
19	110.5	15.5	212287	2	AC108574 Rattus no
20	109.5	15.3	184705	2	AC104910 Mus muscu
21	109.5	15.3	202125	2	AC124636 Mus muscu
22	104.5	14.6	78001	2	AC139696 Homo sapi
23	102.5	14.3	189920	10	AL954353 Mouse DNA
24	102	14.3	278943	2	AC147571 Pan trogl
25	100	14.0	190641	10	AL626775 Mouse DNA
26	100	14.0	286358	2	AL713872 Mus muscu
27	99.5	13.9	139117	2	AC068147 Homo sapi
28	98.5	13.8	34984	3	AC005801 Leishmani
29	98.5	13.8	35573	3	AC009603 Leishmani
30	98.5	13.8	108797	3	AC121231 Leishmani
31	98	13.7	1892	6	BD136402 95 human
32	97.5	13.6	162153	9	AC092331 Homo sapi
33	97.5	13.6	169794	2	AC004688 Plasmodi
34	97.5	13.6	188887	9	AC116364 Homo sapi
35	97.5	13.6	189648	2	AC069163 Homo sapi
36	97.5	13.6	196149	2	AC004709 Plasmodi
37	97.5	13.6	252650	3	AE014847 Plasmodi
38	97	13.6	238748	2	AC120746 Rattus no
39	97	13.6	244341	2	AC108226 Rattus no
40	96.5	13.5	156698	10	AL837509 Mouse DNA
41	96.5	13.5	175340	2	AC110159 Mus muscu
42	96.5	13.5	175754	2	AC110189 Mus muscu
43	96.5	13.5	204081	2	AC130681 Mus muscu
44	96.5	13.5	231930	2	AC134911 Mus muscu
45	96.5	13.5	249025	2	AC131475 Rattus no

ALIGNMENTS

RESULT 1

/protein_id="CAD48808.1"
 /db_xref="GI:23386430"
 /db_xref="RENTREML:CAD48808"
 /translation="MGNVLLRLKSLGGENKPKKGNPDEERKKEQMTTTPERKLODOD
 KKSQSVSSNENSGSSEVCYVINHI:PHORSLSLSSDDGYENIDSLTRKVRQF
 RERSEYFALLATSVSRPCSCTHEHDYEVVFPH"

ORIGIN

Alignment Scores:

Pred. No.: 3,126-65 Length: 3762
 Score: 715.00 Matches: 135
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-10-005-907-2 (1-135) x AX505122 (1-3762)

QY 1 MetGlyAsnTyrluLeuArgLysLeuSerCysLeuGluAsnGlnLysProLys 20
 DB 25 ATGGGAATATCTCTCGGAAACTCAGTTCCTGGGAGAGAAATCAAAAGAGCCCAAG 84
 QY 21 LysGlyAsnProAspGluGluArgLysArgGlnGluMetThrThrPheGluArgLysLeu 40
 DB 85 AAGGAAACCCAGATGAGGAAAGAAACCGCGAGGAATGACTACATTTGAAGAAACTT 144
 QY 41 GlnAspGlnAspLysSerGlnGluValSerSerThrSerAsnGlnGluAsnGluAsn 60
 DB 145 CAAGATCAAGATAAGAAACCCAGAGGTTTCATCCACTTCTAATCAGGAAACGAGAAAT 204
 QY 61 GlySerGlySerGluGluValCysTyThrValIleAsnHisIleProHisGlnArgSer 80
 DB 205 GGCAGTGGTTCGAGAGAGTGTCTACACTGATTAATCATCCCCCATCAGAGATCC 264
 QY 81 SerLeuSerSerAsnAspGlyTyrluGluAsnIleAspSerLeuThrArgLysValArg 100
 DB 265 TCCTCAGCTCCCAATGATGCTATGAGAACATTCCTCCTCACAAGGAAAGTGAGA 324
 QY 101 GlnPheArgLysSerGluThrGluTyrluAlaLeuLeuArgThrSerValSerArgPro 120
 DB 325 CAGTTTAGAGAAAGGTCAGAGACAGAAATATGCCCTCTTAGAGACTCTCTTAGTAGGCCCT 384
 QY 121 CysSerCysThrHisGluHisAspTyrluGluValPheProHis 135
 DB 385 TGTTCTGCACCATGAGCATGATTATGAAGTTGTGTTTCCACAC 429

RESULT 3

AK123798 1709 bp mRNA linear PRI 09-SEP-2003
 LOCUS Homo sapiens cDNA FLJ41804 fis, clone NOV2000710.
 DEFINITION AK123798

VERSION AK123798.1 GI:34529425
 ologo capping; fis (full insert sequence).
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens

ORGANISM

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

AUTHORS

1 Suzuki, O., Sasaki, N., Aotsuka, S., Shoji, T., Ichihara, T.,
 Shiohata, N., Matsumoto, K., Hirano, M., Sano, S., Nomura, R.,
 Yoshikawa, Y., Matsumura, Y., Motiya, S., Chiba, E., Momiyama, H.,
 Onogawa, S., Kaeriyama, S., Satoh, N., Matsumura, H., Takahashi, E.,
 Kataoka, R., Kuga, N., Kuroda, A., Satoh, I., Kamata, K., Takami, S.,
 Teraehina, Y., Watanabe, M., Sugiyama, T., Irie, R., Otsuki, T.,
 Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y.,
 Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H.,
 Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K.,
 Wagatsuma, M., Murakawa, K., Kanehori, K., Takahashi-Fujii, A.,
 Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S.,
 Nagahari, K., Masuno, Y., Nagai, K. and Isogai, T.

NEDO human cDNA sequencing project

TITLE

JOURNAL

Unpublished

REFERENCE

2 (bases 1 to 1709)

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

ORIGIN

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

US-10-005-907-2 (1-135) x AK123798 (1-1709)

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

Isogai, T. and Yamamoto, J.

Direct Submission

Submitted (15-JUL-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7

Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan

(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)

NEDO human cDNA sequencing project supported by Ministry of

Economy, Trade and Industry of Japan; cDNA full insert sequencing;

Research Association for Biotechnology (RAB); cDNA library

construction; Helix Research Institute (HRI) (supported by Japan

Key Technology Center etc.); 5'- & 3'-end one pass sequencing; RAB,

HRI, and Biotechnology Center, National Institute of Technology and

Evaluation; clone selection for full insert sequencing; HRI and

RAB; annotation: HRI and RAB.

Location/Qualifiers

1..1709

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="NOVAR2000710"

/tissue_type="ovary"

/clone_lib="NOVAR2"

/dev_stage="adult"

/note="Cloning vector: pME18SFL3"

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae; Macaca.

1 Osada.N., Hida.M., Kusuda.J., Tanuma.R., Hirata.M., Hirai.M., Terao.K., Suzuki.Y., Sugano.S. and Hashimoto.K.

Prediction of unidentified human genes on the basis of sequence similarity to novel cDNAs from cynomolgus monkey brain

Online Publication

Genome Biology 2001 3(1): research0006.1-0006.5;
http://genomebiology.com/2001/3/1/research/0006/

2 (bases 1 to 1924)
Hashimoto.K., Osada.N., Hida.M., Kusuda.J. and Sugano.S.

Direct Submission
Submitted (27-APR-2001) Katsuyuki Hashimoto, National Institute of Infectious Diseases, Division of Genetic Resources; 23-1, Toyama 1-chome, Shinjuku-ku, Tokyo 162-8640, Japan

(E-mail: khashi@nih.go.jp, URL: http://www.nih.go.jp/yoken/genebank/, Tel: 81-3-5285-1111 (ex. 2120), Fax: 81-3-5285-1181)

Lab host: TOF10

Vector: pME18S-FL3 (Acc.No. AB009864)

R. Site1: DraIII (CACTGTGTG)

R. Site2: DraIII (CACTGTGTG)

Description: 1st strand cDNA was primed with an oligo(dT) primer [ATGTGCGCTTTTCTTTTCTTTT]; double-stranded cDNA was synthesized using specific 5' and 3' primers and amplified by PCR. The PCR product was digested with SfiI and size selection was performed to exclude fragments <1.5kb. The SfiI-digested PCR product was cloned into distinct DraIII sites of pME18S-FL3. XhoI sites just outside the DraIII sites can be used to isolate the cDNA insert. Libraries were constructed by oligo-capping method (Sugano et al., Institute of Medical Science, University of Tokyo).

Custom primer used for sequencing

(5' end primer [CTTCTGCTCTAAAGCTGGG];

3' end primer [CGACCTGAGCTGGAGCA])

Location/Qualifiers

1..1924

/organism="Macaca fascicularis"

/mol_type="mRNA"

/db_xref="taxon:9541"

/clone="Qtra-14007"

/sex="male"

/tissue type="temporal lobe right"

/clone_lib="macaque brain cDNA library Qtra"

/dev_stage="adult"

248..559

/codon_start=1

/product="hypothetical protein"

/protein_id="BAB46903.1"

/db_xref="GI:13874586"

/translation="MTTFERKLQDDKKSEVSIQENENGSGSBEVCVTVINHP

HRSSLSNDDGYENIDSLTRKFRERSSEVALRTSVSRPSTHEDYEVLP

H"

ORIGIN

Alignment Scores:

Pred. No.: 4,95E-57 Length: 1924

Score: 634.00 Matches: 120

Percent Similarity: 96.12% Conservativity: 4

Best Local Similarity: 93.02% Mismatches: 5

Query Match: 88.67% Indels: 0

DB: 9 Gaps: 0

US-10-005-907-2 (1-135) x AB060908 (1-1924)

QY 7 ArgLysLeuSerCysLeuGlyGluAsnGlnLysLysProLysLysGlyAsnProAspGlu 26

170 AGGACATGAGTTCCTGGAGAGATCAAAGAGCCAGGAAGAAACCCAGATGAG 229

QY 27 GluArgLysArgGlnGluMetThrThrPheGluArgLysLeuGlnAspGlnAspLysLys 46

230 GAAAGAAAACGGCAGGAATGACTACATTGTAAGAAAAAATTTCAAGATCAAGATAAGAAA 289

QY 47 SerGlnGluValSerSerThrSerAsnGlnGluAsnGlnAsnGlySerGlySerGluGlu 66

290 AGCAAGAAGATTTCATCCATTCTAATCAGGAACAACGAGAAATGGCAGTGGTCTTGAGAA 349

QY 67 ValCysTyrThrValIleAsnHisIleProHisGlnArgSerSerLeuSerSerAsnAsp 86

350 GTGTGCTACACTGTCTAATTAATCACCATCCCCATGGAGGTCTTCCTGAGTCCAAATGAT 409

QY 87 AspGlyTyrGluAsnIleAspSerLeuThrArgLysValArgGlnPheArgGluArgSer 106

410 GATGCTATGAGAACATTGACTCCCTCAAGGAAGTGGAGAAATTTAGAGAAAGGTCA 469

QY 107 GluThrGluTyrAlaIleuLeuArgThrSerValSerArgProCysSerCysThrHisGlu 126

470 GAGACAGAATATGCGCTCTTAGGACTTCTGTAGTGGCCTTATTCTTGACCCCATGAG 529

QY 127 HisAspTyrGluValValPheProHis 135

530 CATGATTAGGAAGTTGTGCTTCCACAC 556

RESULT 5

AK126682 4032 bp mRNA linear PRI 09-SEP-2003

LOCUS Homo sapiens cDNA FLJ44728 fis, clone BRACE3024537.

DEFINITION AK126682

ACCESSION AK126682

VERSION AK126682.1 GI:34533254

KEYWORDS oligo capping; fis (full insert sequence).

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1

1 Ninomiya,K., Wagatsuma,M., Kanda,K., Kondo,H., Yokoi,T.,

Kodaira,H., Furiya,T., Takahashi,N., Kikkawa,E., Omura,Y., Abe,K.,

Kamihara,K., Katsuta,N., Sato,K., Tanikawa,M., Yamazaki,M.,

Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S.,

Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T.,

Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M.,

Kikuchi,H., Murakawa,K., Kanehori,K., Takahashi-Fujii,A.,

Oshina,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S.,

Nagahari,K., Masubo,Y., Nagai,K. and Isogai,T.

NEDO human cDNA sequencing project

Unpublished

REFERENCE 2 (bases 1 to 4032)

Isogai,T. and Yamamoto,J.

Direct Submission

Submitted (15-JUL-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7

Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan

(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)

NEDO human cDNA sequencing project supported by Ministry of

Economy, Trade and Industry of Japan; cDNA full insert sequencing:

Research Association for Biotechnology (RAB); cDNA library

construction: Helix Research Institute (HRI) (supported by Japan

Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,

HRI, and Biotechnology Center, National Institute of Technology and

Evaluation; clone selection for full insert sequencing: HRI and

RAB; annotation: HRI and RAB.

Location/Qualifiers

1..4032

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="BRACE3024537"

/tissue_type="cerebellum"

/clone_lib="BRACE3"

/notes="cloning vector: pME18SFL3"

FEATURES

source

ORIGIN

Alignment Scores:

Pred. No.: 2.02E-51 Length: 4032

Score: 584.00 Matches: 115

Percent Similarity: 85.19% Conservativity: 0

Best Local Similarity: 85.19% Mismatches: 0
 Query Match: 81.68% Indels: 20
 DB: 9 Gaps: 1

US-10-005-907-2 (1-135) x AK126682 (1-4032)

QY 1 MetGlyAsnTyrLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLysProLys 20
 DB 377 ATGGGAATATCTCTCGCAAACTC----- 403
 QY 21 LysGlyAsnProAspGluGluArgLysArgGlnGluMetThrThrPheGluArgLysLeu 40
 DB 404 -----AGCAGGAATGACTACATTTGAGAAACTT 436
 QY 41 GlnAspGlnAspLysSerGlnGluValSerSerThrSerAsnGlnGluAsn 60
 DB 437 CAAGATCAAGATAAGAAAGCAAGAGTTTCACATCTTAATCAGGAAACGAGAAT 496
 QY 61 GlySerGlySerGluGluValCysThrValIleAsnHisIleProHisGlnArgSer 80
 DB 497 GGCAGTGGTCTTGAGAGGTGCTGACACTGTANTATCATATCCATCCCATCAGATCC 556
 QY 81 SerLeuSerSerAsnAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArg 100
 DB 557 TCCTGAGTCCATGATGATGGCTATGAGACATGACTCCCTCAAGAAAGTGAGA 616
 QY 101 GlnPheArgLysGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArgPro 120
 DB 617 CAGTTTAGAAGAAAGGTGAGAGACAGATATGCCCTCTTAGGACTTCTGTAGAGCCT 676
 QY 121 CysSerCysThrHisGluHisAspTyrGluValValPheProHis 135
 DB 677 TGTTCCTGCACCATGAGCATGATTATGAAGTTGTGTTCACAC 721

RESULT 6

AK124520 4136 bp mRNA linear PRI 09-SEP-2003
 LOCUS Homo sapiens cDNA FLJ42529 fis, clone BRACE3002390.
 DEFINITION AK124520

ACCESSION AK124520.1 GI:34530324

VERSION oligo capping, fis (full insert sequence).

KEYWORDS Homo sapiens (human)

SOURCE

Organism Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 Ninomiya K., Wagatsuma M., Kanda K., Kondo H., Yokoi T.,
 Kodaira H., Furiya T., Takahashi M., Kikkawa E., Omura Y., Abe K.,
 Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,
 Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S.,
 Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
 Kimura K., Yamashita H., Matsuo K., Nakamura Y., Sekine M.,
 Kikuchi H., Murakawa K., Kanehori K., Takahashi-Fujii A.,
 Oshima A., Sugiyama A., Kawakami B., Suzuki Y., Sugano S.,
 Nagahara K., Maeho Y., Nagai K. and Isogai T.
 NEDO human cDNA sequencing project

TITLE

Unpublished

JOURNAL

2 (bases 1 to 4136)

REFERENCE

Isogai T. and Yamamoto J.

AUTHORS

Direct Submission

TITLE

Submitted (15-JUL-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7
 Kazusa-Kanagari, Kisarazu, Chiba 292-0818, Japan
 (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
 NEDO human cDNA sequencing project supported by Ministry of
 Economy, Trade and Industry of Japan; cDNA full insert sequencing:
 Research Association for Biotechnology (RAB); cDNA library
 construction: Helix Research Institute (HRI) (supported by Japan
 Key Technology Center etc.); 5' & 3'-end one pass sequencing: RAB,
 HRI, and Biotechnology Center, National Institute of Technology and
 Evaluation; clone selection for full insert sequencing: HRI and
 RAB; annotation: HRI and RAB.

FEATURES

Location/Qualifiers

1..4136

source

/organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="BRACE3002390"
 /issue_type="cerebellum"
 /clone_lib="BRACE3"
 /note="Cloning vector: pME18SFL3"
 136..537
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="BAC85873.1"
 /db_xref="GI:34530325"
 /translation="MVMLEDNSNTGCGARNCECLVFLSVLGCQSRKGMRTQQAG
 RWLRAGREASSETNPETGLILDQSPFLPLAARGWQHEPVRVNCVLPHPSSNCFQG
 NPFKWSQTQLNRVTVKRWELIISSENSGRK"

CDS

ORIGIN

Alignment Scores: 4.3e-51 Length: 4136
 Pred. No.: 581.00 Matches: 114
 Score: 581.00 Conservative: 1
 Percent Similarity: 85.19% Mismatches: 0
 Best Local Similarity: 84.44% Indels: 20
 Query Match: 81.26% Gaps: 1
 DB: 9
 US-10-005-907-2 (1-135) x AK124520 (1-4136)
 QY 1 MetGlyAsnTyrLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLysProLys 20
 DB 498 ATGGGAATATCTCTCGCAAACTC----- 524
 QY 21 LysGlyAsnProAspGluGluArgLysArgGlnGluMetThrThrPheGluArgLysLeu 40
 DB 525 -----AGCAGGAATGACTACATTTGAGAAACTT 557
 QY 41 GlnAspGlnAspLysSerGlnGluValSerSerThrSerAsnGlnGluAsnGluAsn 60
 DB 558 CAAGATCAAGATAAGAAAGCAAGAGTTTCACATCTTAATCAGGAAACGAGAAT 617
 QY 61 GlySerGlySerGluGluValCysThrValIleAsnHisIleProHisGlnArgSer 80
 DB 618 GGCAGTGGTCTTGAGAGGTGCTGACACTGTCAATATCATATCCATCCCATCAGAAATCC 677
 QY 81 SerLeuSerSerAsnAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArg 100
 DB 678 TCCTGAGTCCATGATGATGGCTATGAGACATGACTCCCTCAAGAAAGTGAGA 737
 QY 101 GlnPheArgLysGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArgPro 120
 DB 738 CAGTTTAGAAGAAAGGTGAGAGACAGATATGCCCTCTTAGGACTTCTGTAGAGCCT 797
 QY 121 CysSerCysThrHisGluHisAspTyrGluValValPheProHis 135
 DB 798 TGTTCCTGCACCATGAGCATGATTATGAAGTTGTGTTCACAC 842

RESULT 7

AC074365/c

LOCUS

DEFINITION

AC074365

ACCESSION

AC074365.5

VERSION

HTG; HTGS_PHASE1; HTGS_DRAFT.

KEYWORDS

Homo sapiens (human)

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 141268)

AUTHORS

The sequence of Homo sapiens clone

Unpublished

JOURNAL

2 (bases 1 to 141268)

REFERENCE

AUTHORS

Waterston, R.H.

AC074365 141268 bp DNA linear HTG 23-SEP-2000
 Homo sapiens chromosome 1 clone RF11-115C4, WORKING DRAFT SEQUENCE,
 10 unordered pieces.

AC074365

AC074365.5

HTG; HTGS_PHASE1; HTGS_DRAFT.

KEYWORDS

Homo sapiens (human)

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 141268)

AUTHORS

The sequence of Homo sapiens clone

Unpublished

JOURNAL

2 (bases 1 to 141268)

REFERENCE

AUTHORS

Waterston, R.H.

TITLE Direct Submission
 JOURNAL Submitted (29-JUL-2000) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 COMMENT On Sep 23, 2000 this sequence version replaced gi:9838075.

----- Genome Center -----
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: http://genome.wustl.edu/gsc/index.shtml
 ----- Project Information -----
 Center project name: H.NH0115C04
 ----- Summary Statistics -----
 Sequencing vector: M13; 100%
 Sequencing method: plasmid; 0%
 Chemistry: Dye-terminator Big Dye; 0% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 136274 bases at least Q40
 Consensus quality: 138356 bases at least Q30
 Consensus quality: 139182 bases at least Q20
 Insert size: 142000; agarose-1p
 Insert size: 140368; sum-of-contigs
 Quality coverage: 4.64 in Q20 bases; agarose-1p
 Quality coverage: 4.81 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 10 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

* 1 31377: contig of 31377 bp in length
 * 31378 31477: gap of unknown length
 * 31478 63011: contig of 31534 bp in length
 * 63012 63111: gap of unknown length
 * 63112 64845: contig of 1734 bp in length
 * 64846 67606: gap of unknown length
 * 67607 67706: contig of 2661 bp in length
 * 67707 67774: gap of unknown length
 * 67775 77774: contig of 9968 bp in length
 * 77775 89256: contig of 11482 bp in length
 * 89257 89356: gap of unknown length
 * 89357 100401: contig of 11044 bp in length
 * 100401 100500: gap of unknown length
 * 100501 111915: contig of 11415 bp in length
 * 111916 125713: gap of unknown length
 * 125713 125814: contig of 13698 bp in length
 * 125814 141268: contig of 15455 bp in length.

FEATURES
 source 1. 141268
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="1"
 /clone="RP11-115C4"
 1. 31377
 /note="assembly_name:Contig10"
 31478..63011
 /note="assembly_name:Contig11"
 63112..64845
 /note="assembly_name:Contig2"
 64946..67606
 /note="assembly_name:Contig3"
 67707..77674
 /note="assembly_name:Contig4"
 77775..89256
 /note="assembly_name:Contig5"
 89357..100400

misc_feature
 misc_feature
 misc_feature
 misc_feature
 misc_feature
 misc_feature
 misc_feature

/note="assembly_name:Contig6"
 100501..111915
 /note="assembly_name:Contig7"
 112016..125713
 /note="assembly_name:Contig8"
 125814..141268
 /note="assembly_name:Contig9"
 clone_end:SP6
 vector_side:left

misc_feature
 misc_feature
 misc_feature

ORIGIN
 Alignment Scores:
 Pred. No.: 1.58e-33 Length: 141268
 Score: 430.00 Matches: 80
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 60.14% Indels: 0
 DB: 2 Gaps: 0
 US-10-005-907-2 (1-135) x AC074365 (1-141268)
 QY 56 GlnGluAsnGluAsnGlySerGlySerGluGluValCysTyrThrValIleAsnHisIle 75
 DB 5700 CAGGAACAGAAAGGCGAGTGGTCTTCTGAAGAAGTGTGCTACACTGTCTATTATCACATC 5641
 QY 76 ProHisGlnArgSerSerLeuSerSerAsnAspGlyTyrGluAsnIleAspSerIleu 95
 DB 5640 CCCATCAGAGATCCTCCCTGAGCTCCATGATGCTATGAGATTCATTCACCTC 5581
 QY 96 ThrArgLysValArgGlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThr 115
 DB 5580 ACAAGAAAGTGAAGACAGTTTAGAAGAGGTCAGACAGAGATATGCCCTCTTTAGGACT 5521
 QY 116 SerValSerArgProCysSerCysThrHisGluHisAspTyrGluValValPheProHis 135
 DB 5520 TCTGTAGTAGGCCTGTCTCTGCACCATGAGCATGATTATGAGTGTGTTCACAC 5461

RESULT 8
 LOCUS AL606804
 DEFINITION Human DNA sequence from clone RP11-978115 on chromosome 1, complete sequence.
 ACCESSION AL606804
 VERSION AL606804.11 GI:24366459
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 185467)
 Almeida,J.
 Direct Submission
 Submitted (16-OCT-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
 On Oct 24, 2002 this sequence version replaced gi:18121511.
 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr1>. RP11-978115 is from the library RP11-11.4 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pBACe3.6

----- Genome Center

Center: Wellcome Trust Sanger Institute

Center code: SC

Web site: <http://www.sanger.ac.uk>

Contact: humquery@sanger.ac.uk

FEATURES

source
1. 185467
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP11-978115"
/clone_lib="RP11-11.4"

ORIGIN

Alignment Scores:
Pred. No.: 2.13e-33 Length: 185467
Score: 430.00 Matches: 80
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 60.14% Indels: 0
DB: 9 Gaps: 0

US-10-005-907-2 (1-135) x AL606804 (1-185467)

Qy 56 GlnGluAsnGluAsnGlySerGlySerGluGluValCysTyrThrValIleAsnHisIle 75
Db 132080 CAGGAAACAGGAATGCGAGTGGTTCGGAAGAGTGTGTACACTGTCAATATCATCATC 132139
Qy 76 ProHisGlnArgSerLeuSerSerAsnAspGlyTyrGluAsnIleAspSerLeu 95
Db 132140 CCCCATCAGAGATCTCCCTCGAGTCCAAATGATGCGTATGAGACATGACTCCCTC 132199
Qy 96 ThrArgIysValArgGlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThr 115
Db 132200 ACAAGAAAGTGAACAGATTTAGGAAGGTGACAGACAGAAATGCGCTTCTTAGACT 132259
Qy 116 SerValSerArgProCysSerCysThrHisGluHisAspTyrGluValValPheProHis 135
Db 132260 TCTCTTAGTAGGCTTGTCTGTCACCCATGAGCATGATTATGAAGTTGTGTTTCCACAC 132319

RESULT 9
BC030506 1532 bp mRNA linear PRI 06-OCT-2003
LOCUS Homo sapiens germinal center expressed transcript 2, mRNA (CDNA
DEFINITION clone MGC:40441 IMAGE:4385178), complete cds.
ACCESSION BC030506
VERSION BC030506
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1532)
AUTHORS Klausner R.D., Feingold E.A., Grouse L.H., Derge J.G., Strausberg R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J., Bosak S.A., McEwan P.J.,

McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalios, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
12477932
2 (bases 1 to 1532)
Strausberg, R.
Direct Submission
Submitted (07-MAY-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: Lou Staudt
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: the I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc.mgc@nih.gov
Akhter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Hagnigni, P., Hansen, N., Ho, S.-L., Karlins, S., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantrop, S., Thomas, P.J., Touchman, J.W., Tsurgon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAP Plate: 64 Row: a Column: 15
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 22749536.

FEATURES

source

1. 1532
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:40441 IMAGE:4385178"
/tissue_type="Lymph, lymphoma"
/clone_lib="NIH MGC_85"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1. 1532
/gene="GCET2"
/note="synonyms: HGAL, GCAT2, MGC40441"
/db_xref="LocusID:257144"
/db_xref="MIM:607792"
54..590
/codon_start=1
/product="germinal center expressed transcript 2"
/protein_id="AAH30506.1"
/db_xref="GI:20987806"
/translation="MGSLLENRRQQTQEMPNVNRMQSPKORTSCWDHHAEGCF
CLPWKLLIFPKQDSQNERMSSTPIQDNVDQTYSELCTYLNEHVLCTRPSGNS
AEYEYVPCAAKRPRESLGTETETSLHMPSTDFRARSFEDEYELLMFRRISSHF
LQQPRLMAPSETQFSLH"

gene

CDS

ORIGIN

Alignment Scores:
 Pred. No.: 2,76e-08 Length: 1532
 Score: 169.50 Matches: 47
 Percent Similarity: 51.61% Conservative: 33
 Best Local Similarity: 30.32% Mismatches: 54
 Query Match: 23.71% Indels: 21
 DB: 9 Gaps: 5

US-10-005-907-2 (1-135) x BC030506 (1-1532)

QY 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLysPro--- 19
 DB 54 ATGGGAATTCCTCTCTGAGAGAAACAGCGCGCAGACACTCAAGAGATCGCTGG 113
 QY 20 -----LysLysGlyAsnProAspGluGluArgLysLysArg----- 30
 DB 114 AATGTGAGATGCAAAAGCCCAACAGAGAAACATCCAGATCTGGATCACCATATCGCT 173
 QY 31 -----GlnGluMetThrThrPheGluArgLysLeuGluAsp 42
 DB 174 GAAGGGTGTTCCTGCTCCATGGAAAAAATACTCATTTTGAAGAGG---CAAGAT 230
 QY 43 GlnAspLysLysSerGlnGluValSerThrSerAsnGlnGluAsnGlnGlySer 62
 DB 231 TCCCAAAACGAAATCAAGAGTACTATCCATCCAGACCAATGTTGACAGACC 290
 QY 63 GlySerGluGluValCysThrValIleAsnHis---IleProHisGlnArgSerSer 81
 DB 291 TACTCAGAGGAGTGTCTATACCTCATCAATCATCGGTCTCTGTACAGGCCATCA 350
 QY 82 LeuSerAsnAspAspGlyTyrGluAsnLeuAspSerLeuThrArgLysValArgGln 101
 DB 351 GGGAACTCTGCTGAAGAGTACTATGAGAATGTTCCCTGCAAGCTGAGAGACCCAGAG 410
 QY 102 PheArgGluArgSerGluThrValAlaLeuArgLysValArgLysValArgGln 121
 DB 411 TCCCTGGAGGAGTGTCTATACCTCATCAATCATCGGTCTCTGTACAGGCCATCA 470
 QY 122 SerCys---ThrHisGluHisAspTyrGluValValPheProHis 135
 DB 471 CATGCCCGATCCCGAGAGATGAATGAATGAACTTCTCATGCCCTCAC 515

RESULT 10
 AF521911 1654 bp mRNA linear PRI 14-JAN-2003
 LOCUS
 DEFINITION Homo sapiens HGAL mRNA, complete cds.
 ACCESSION AF521911
 VERSION AF521911.1 GI:27733682
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 LOSSOS I.S., Alizadeh, A., Rajapaksa, R., Tibshirani, R. and Levy, R.
 HGAL is a novel interleukin-4-inducible gene that strongly predicts survival in diffuse large B-cell lymphoma
 Blood 101 (2), 433-440 (2003)
 JOURNAL
 MEDLINE 22397484
 PUBMED 12509382
 REFERENCE 2 (bases 1 to 1654)
 AUTHORS Lossos, I.S., Alizadeh, A.A. and Levy, R.
 TITLE Direct Submission
 JOURNAL Submitted (16-JUN-2002) Division of Oncology, Department of Medicine, Stanford University, 269 Campus Drive, CCSR building, Room 1100, Palo Alto, CA 94305-5151, USA
 FEATURES
 source
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 173..709
 /note="germinal center-associated lymphoma; contains

modified ITAM motif, similar to mouse M17"
 /codon_start=1
 /product="HGAL"
 /protein_id="AA021701.1"
 /db_xref="GI:27733683"
 /translation="MNSLLRENRQNTQEMVNRMQSPKORTSRQDHHIARGCF
 CLPWEKILIFKQDSQENERSMSTPIQDAVDQYSELVCLINHLVLCITRPSGNS
 AEYENVPKAEPRRESLGGTETYSILHMPSTDFRARSPEDEYELMHPRISSH
 LQOPRLMAPSETQFSLH"

ORIGIN
 Alignment Scores:
 Pred. No.: 3e-08 Length: 1654
 Score: 169.50 Matches: 47
 Percent Similarity: 51.61% Conservative: 33
 Best Local Similarity: 30.32% Mismatches: 54
 Query Match: 23.71% Indels: 21
 DB: 9 Gaps: 5

US-10-005-907-2 (1-135) x AF521911 (1-1654)

QY 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLysPro--- 19
 DB 173 ATGGGAATTCCTCTCTGAGAGAAACAGCGCGCAGACACTCAAGAGATCGCTGG 232
 QY 20 -----LysLysGlyAsnProAspGluGluArgLysLysArg----- 30
 DB 233 AATGTGAGATGCAAAAGCCCAACAGAGAAACATCCAGATCTGGATCACCATATCGCT 292
 QY 31 -----GlnGluMetThrThrPheGluArgLysLeuGluAsp 42
 DB 293 GAAGGGTGTTCCTGCTCCATGGAAAAAATACTCATTTTGAAGAGG---CAAGAT 349
 QY 43 GlnAspLysLysSerGlnGluValSerThrSerAsnGlnGluAsnGlnGlySer 62
 DB 350 TCCCAAAACGAAATCAAGAGTACTATCCATCCAGACCAATGTTGACAGACC 409
 QY 63 GlySerGluGluValCysThrValIleAsnHis---IleProHisGlnArgSerSer 81
 DB 410 TACTCAGAGGAGTGTCTATACCTCATCAATCATCGGTCTCTGTACAGGCCATCA 469
 QY 82 LeuSerAsnAspAspGlyTyrGluAsnLeuAspSerLeuThrArgLysValArgGln 101
 DB 470 GGAATCTGTCTGAAGAGTACTATGAGAATGTTCCCTGCAAGCTGAGAGACCCAGAG 529
 QY 102 PheArgGluArgSerGluThrValAlaLeuArgLysValArgLysValArgProCys 121
 DB 530 TCCTTGGAGGAACTCAGACTGAGTATTCATCTTACATATGCTTCTACAGACCCAGG 589
 QY 122 SerCys---ThrHisGluHisAspTyrGluValValPheProHis 135
 DB 590 CATGCCCGATCCCGAGAGATGAATGAATGAACTTCTCATGCCCTCAC 634

RESULT 11
 AF521246 3270 bp mRNA linear PRI 14-OCT-2003
 LOCUS
 DEFINITION Homo sapiens germinal center B-cell expressed transcript 2 (GCET2)
 mRNA, complete cds.
 ACCESSION AF521246
 VERSION AF521246.1 GI:27948576
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 PAN, Z., Shen, Y., Du, C., Zhou, G., Rosenwald, A., Staudt, L.M., Greiner, T.C., McKeithan, T.W. and Chan, W.C.
 Two newly characterized germinal center B-cell-associated genes, GCET1 and GCET2, have differential expression in normal and neoplastic B cells
 Am. J. Pathol. 163 (1), 135-144 (2003)
 JOURNAL 22702315
 MEDLINE

PUBMED 12819018
REFERENCE 2 (bases 1 to 3270)
AUTHORS Pan,Z., Shen,Y., Du,C., Zhou,G., McKeithan,T., Rosenwald,A., Staudt,J., and Chan,W.C.
TITLE Direct Submission
JOURNAL Submitted (07-JAN-2003) Pathology, University of Nebraska Medical Center, #42, Omaha, NE 68198, USA
FEATURES Location/Qualifiers
source 1..3270
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="3"
/map="3q13.13"
/cell_line="DHL16"
gene 1..3270
/gene="GCBT2"
/notes="GCBT2"
/notes="GCBT2"
107..643
/gene="GCBT2"
/notes="GCBT2"
CDS
/notes="similar to Mus musculus germinal center transcript M17; similar to MGC40441 sequence in GenBank Accession. Number BC030506"
/codon_start=1
/product="germinal center B-cell expressed transcript 2"
/protein_id="AA022147.1"
/db_xref="GI:27948577"
/translation="MGNLSLRNRROQNTQEMPNVRMQSPKORTSRCDWHIABGCF CLPWKILIEKQDQSONENRMSSTPIQNDVDTYSEELCYTLINRVLCTRPQSGN AEEYVNPCKARPRPSLGTTETYSLLHMPSTDPRHARSPEDEYELLMPHRISSH LQPRPLMAPSERQFSLH"
ORIGIN
Alignment Scores:
Pred. No.: 6,34e-08 Length: 3270
Score: 169.50 Matches: 47
Percent Similarity: 51.61% Conservatives: 33
Best Local Similarity: 30.32% Mismatches: 54
Query Match: 23.71% Indels: 21
DB: 9 Gaps: 5
US-10-005-907-2 (1-135) x AY212246 (1-3270)
QY 1 MetGlyAsnTyrLeuArgGlyLeuSerCysLeuGlyGluAsnGlnLysPro--- 19
Db 107 ATGGAAATTCCTGCTGAGAGAAACAGCGGCGAGCAGACACTCAAGAGATGCTTGG 166
QY 20 -----LysLysGlyAsnProAspGluArgLysArg----- 30
Db 167 AACGTGAGATGCAAGAGCCCAACAGAGAACATCCAGATGCTGGGATCACCATATCGCT 226
QY 31 -----GlnGluMetThrThrPheGluArgGlyLeuGlnAsp 42
Db 227 GAAGGTGTTCTGCTTCCATGCGAAATAATCTCATTTTGAAGAGAG---CAAGAT 283
QY 43 GlnAspLysSerGlnGluValSerThrSerAsnGlnGluAsnGlySer 62
Db 284 TCCCAAAACGAAATGAAGAATGTCATCTACTCCATCCAGGACAATGTTGACCAACC 343
QY 63 GlySerGluGluValCysTyrThrValIleAsnHis---IleProHisGlnArgSer 81
Db 344 TACTCAGAGGAGCTGTGCTACCCCTCATCAATCATCGGTTCTCTGTACAGGCCATCA 403
QY 82 LeuSerSerAsnAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArgGln 101
Db 404 GGGAACTCTGCTGAGAGTACTATGAGATGTTCTCCGCAAGCTGAGAGCCCAAGAG 463
QY 102 PheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArgProCys 121
Db 464 TCCITGGGAGGAACCTGAGACTGAGTATTCACTTCTTACATATGCTTCTTACAGCCCAGG 523
QY 122 SerCys---ThrHisGluHisAspTyrGluValValPheProHis 135

Db 524 CATGCCCGCATCCCGAAGATGAATATGAATCTTCTCATGCTCCAC 568
RESULT 12
AX740457
LOCUS AX740457 1643 bp DNA linear PAT 10-MAY-2003
DEFINITION Sequence 46 from Patent WO02020756.
ACCESSION AX740457
VERSION AX740457.1 GI:30523621
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Stuart,J., Lincoln,S.E., Altus,C.M., Dufour,G., Chalup,M., Hillman,J.L., Jones,A.L., Yu,J.Y., Wright,R.J., Gietzen,D., Liu,T., Yap,P., Dahl,C.R., Momiya,M.G., Bradley,D., Rohatgi,S., Harris,B., Roseberry, Ann.M., Gerstin,E.H., Peralta,C.H., David M.H., Panzer, Scott.R., Flores,V., Daffo,A., Marwaha,R., Chen,A.J., Chang,S.C., Au,A.P. and Inman,R.R.
TITLE Secretory molecules
JOURNAL Patent: WO 02020756-A 46 14-MAR-2002;
INCYTE Genomics, Inc. (US)
FEATURES Location/Qualifiers
source 1..1643
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: LG:018739.2:2000SEP08"
ORIGIN
Alignment Scores:
Pred. No.: 3,79e-07 Length: 1643
Score: 159.00 Matches: 47
Percent Similarity: 50.64% Conservatives: 22
Best Local Similarity: 30.13% Mismatches: 55
Query Match: 22.24% Indels: 22
DB: 6 Gaps: 6
US-10-005-907-2 (1-135) x AX740457 (1-1643)
QY 1 MetGlyAsnTyrLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysPro--- 19
Db 1096 ATGGGAAATTCCTGCTGAGAGAAACAGCGGCGAGCAGACACTCAAGAGATGCTTGG 1155
QY 20 -----LysLysGlyAsnProAspGluArgLysArg----- 30
Db 1156 ATGTGAGATGCAAGAGCCCAACAGAGAACATCCAGATGCTGGGATCACCATATCGCT 1215
QY 31 -----GlnGluMetThrThrPheGluArgLysLeuGlnAsp 42
Db 1216 GAAGGTGTTCTGCTTCCATGCGAAATAATCTCATTTTGAAGAGAG---CAAGAT 1272
QY 43 GlnAspLysLysSerGlnGluValSerThrSer---AsnGlnGluAsnGlnGly 61
Db 1273 TCCCAAAACGAAATGAAGAATGTCATCTACTCCATCCAGCAGACATGTTGACCAG 1332
QY 62 SerGlySerGluGluValCysTyrThrValIleAsnHis---IleProHisGlnArgSer 80
Db 1333 ACCTACTCAGAGAGCTGTGCTATACCCCTCATCAATCATCGGTTCTCTGTACAAGGCCA 1392
QY 81 SerLeuSerSerAsnAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArg 100
Db 1393 TCAGGAATCTGCTGAGAGTACTATGAGATGTTCCCTCGCAAGCTGAGAGCCCA 1452
QY 101 GlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArgPro 120
Db 1453 GAGTCTCTGGGAGGAACCTGAGACTGAGTATTCACTTCTTANATATGCTTCTTACAGACCCC 1512
QY 121 CysSerCys---ThrHisGluHisAspTyrGluValValPheProHis 135
Db 1513 AGGCATGCCCGATCCCGCAGAGATGAATATGAATCTTCTCATGCTCCAC 1560

```
RESULT 13
MMU13263
LOCUS Mus musculus M17 mRNA, complete cds. 1672 bp mRNA linear ROD 15-APR-1996
DEFINITION Mus musculus M17 mRNA, complete cds.
ACCESSION U13263
VERSION U13263.1 GI:533130
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 1672)
AUTHORS Christoph, T., Rickert, R. and Rajewsky, K.
TITLE M17: a novel gene expressed in germinal centers
JOURNAL Int. Immunol. 6 (8), 1203-1211 (1994)
PUBMED 7981148
REFERENCE 2 (bases 1 to 1672)
AUTHORS Rickert, R.C.
TITLE Direct Submission
JOURNAL Submitted (11-AUG-1994) Robert C. Rickert, Institute for Genetics, University of Cologne, Weyertal 121, Cologne, 50823, Germany
FEATURES
source
1. 1672
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
gene
1. 1672
/gene="M17"
CDS
94..573
/gene="M17"
/codon_start=1
/product="M17 protein"
/protein_id="AA096821.1"
/db_xref="GI:1262852"
/translation="MGNCLQRTYFRGWSGCCSVSCCLPWNKIRTFKARQESPKQ
NEGTSAPVDNANETTELCLYLDVHEAVRPSVNPAGFYENISKAERHKSS
RGTEYSLVLRFPSPQPLPSTDDYELLPSRFSHAFQPRPLTPYTHFSYQ"
585..727
/gene="M17"
3'UTR
/gene="M17"
ORIGIN
Alignment Scores:
Pred. No.: 0.000703 Length: 1672
Score: 128.00 Matches: 31
Percent Similarity: 59.43% Conservative: 32
Best Local Similarity: 29.25% Mismatches: 39
Query Match: 17.90% Indels: 4
DB: 10 Gaps: 4
US-10-005-907-2 (1-135) x MMU13263 (1-1672)
QY 31 GlnGluMetThrThrPheGluArgLysLeuGlnAspGlnAspLysSerGlnGluVal 50
DB 181 AABAACATACGCACATTAAAGCCAGA---CAAGAGTCTCCAAAGCAAAATGAGGATG 237
QY 51 SerSerThrSerAsnGlnGluAsnGlySerGlySerGluGluValCysTyThr 70
DB 238 ACTTCAGCTCCGTTTCAGGACAATGTAATGAGACCTACAGAGAGTGTGTCTATC 297
QY 71 ValIleAsnHis---IleProHisGlnArgSerSerLeuSerSerAsnAspGlyTyr 89
DB 298 CTGTGTGATCAGAGCTGTGAGGAGAGCCATCAGTGAACCTGTGAGGGTCTTAC 357
QY 90 GluAsnIleAspSerLeuThrArgLysValArgGlnPheArgGluArgSerGluThrGlu 109
DB 358 GAGAACATCTCTTAAACAGCTGAGACACACAAAGAGCTCTTCAAGAGGAAACAGACTGAG 417
QY 110 TyrAlaLeuLeuArg---ThrSerValSerArgProCysSerCysThrHisGluHisAsp 128
DB 418 TATTCGGTTCCTCCGTTTCCTCTCCCTCAGCCCTACCTTCCACA---GATGATGAA 474
```

```
129 TyrGluValValPhePro 134
475 TATGAACCTCTTATGCCC 492
RESULT 14
AK125521/c
LOCUS Homo sapiens CDNA FLJ43533 fis, clone PLACE7001544.
DEFINITION Homo sapiens CDNA FLJ43533 fis, clone PLACE7001544.
ACCESSION AK125521
VERSION AK125521.1 GI:34531645
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Tashiro, H., Yamazaki, M., Watanabe, K., Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Wagatsuma, M., Murakawa, K., Kanehori, K., Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahari, K., Masuho, Y., Nagai, K. and Isogai, T.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 4218)
AUTHORS Isogai, T. and Yamamoto, J.
TITLE Direct Submission
JOURNAL Submitted (15-JUL-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
COMMENT NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5' & 3' end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.
FEATURES
Location/Qualifiers
1..4218
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PLACE7001544"
/tissue_type="placenta"
/clone_lib="PLACE7"
/note="cloning vector: pME18SFL3"
ORIGIN
Alignment Scores:
Pred. No.: 0.00315 Length: 4218
Score: 126.00 Matches: 30
Percent Similarity: 59.76% Conservative: 19
Best Local Similarity: 36.59% Mismatches: 31
Query Match: 17.62% Indels: 2
DB: 9 Gaps: 2
US-10-005-907-2 (1-135) x AK125521 (1-4218)
QY 56 GlnGluAsnGluAsnGlySerGlySerGluGluValCysTyThrValIleAsnHis--- 74
DB 3384 CAGGACAATGTTGACACACCTACTCAGAGGAGTGTGTCTATCCTCATCATCGG 3325
QY 75 IleProHisGlnArgSerSerLeuSerSerAsnAspGlyTyGluAsnIleAspSer 94
DB 3324 GTTCTCTGTACAGGCCCATCAGGGAACCTGCTGAGAGTACTATGAGAAUUTTCCTGCG 3265
QY 95 LeuThrArgLysValArgGlnPheArgGluArgSerGluThrGluTyAlaLeuLeuArg 114
DB 3264 AAGCTGAGAGAGCCAGAGAGTCTTGGAGGAGGAGTCTGAGTATTCTTCTTCTCAT 3205
```

QY 115 ThrSerValSerArgProCysCys---ThrHisGluHisAspTyrGluValValPhe 133
 Db 3204 ATGCTTCTACAGACCCAGGATGCCCCGATCCGAGAGATGATATGACTTCTCATG 3145
 QY 134 ProHis 135
 Db 3144 CCTCAC 3139
 RESULT 15
 AC128688
 LOCUS Homo sapiens 3 BAC RP11-757F18 (Roswell Park Cancer Institute Human
 DEFINITION BAC Library) complete sequence.
 AC128688
 VERSION AC128688.4 GI:22549624
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Eukaryota; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 REFERENCE 1 (bases 1 to 81001)
 AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
 Alsbrooks,S.L., Amarantunge,H.C., Are,J.R., Ayele,M., Banks,T.,
 Barbara,J., Benton,J., Bimge,K., Blankenburg,K., Bonnin,D.,
 Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,N., Bryant,N.P.,
 Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
 Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
 Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
 Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
 Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
 Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
 Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
 Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
 Falle,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
 Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
 Gorrill,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
 Harris,C., Harris,K., Hart,M., Havlik,P., Hawes,A., He,X.,
 Hernandez,J., Hernandez,O., Hodgson,A., Hognes,M., Holloway,C.,
 Hollins,B., Honsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
 Jackson,L.E., Jacobson,B., Jia,X., Johnson,R., Jolivet,S.,
 Joudan,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
 Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
 Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
 Loughes,H., Lozada,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
 Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
 Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M.,
 Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
 Moore,S., Morgan,M., Moorish,T., Morris,S., Moser,M., Neal,D.,
 Nelson,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N.,
 Nickerson,E., Nwokenwo,S., Oguh,M., Okwuonu,G., Oragunye,N.,
 Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L.,
 Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M.,
 Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Saverly,G., Scherer,S.,
 Scott,G., Shen,H., Shoshitari,N., Sisson,I., Sodergren,E.,
 Sonake,F., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A.,
 Taylor,P., Telford,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C.,
 Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L.,
 Vera,V., Villalón,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S.,
 Warren,R., Washington,C., Watlington,S., Williams,G.,
 Williamson,A., Wlezyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y.,
 Wu,Y.F., Zhou,J., Zorrilla,S., Naylor,S.L., Weinstein,G. and
 Gibbs,R.
 TITLE Direct Submission
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 81001)
 AUTHORS Worley,K.C.
 TITLE Direct Submission
 JOURNAL Submitted (22-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 REFERENCE 3 (bases 1 to 81001)
 AUTHORS Worley,K.C.
 TITLE Direct Submission

JOURNAL

REFERENCE
 AUTHORS
 TITLE
 JOURNAL

COMMENT

Submitted (26-AUG-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 4 (bases 1 to 81001)
 Worley,K.C.
 Direct Submission
 Submitted (30-AUG-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Aug 30, 2002 this sequence version replaced gi:22475264.
 INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email
 gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the
 entire insert of this clone. Overlapping regions of clones are only
 sequenced and submitted once, so the sequence for the remainder of
 the insert may be found in the record for the adjacent clones.
 Overlapping clones are noted at the beginning and end of the
 Features listing.

ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches
 of a local database that includes entries from dbSTS, GDB, and
 local mapping efforts.
 Repeats are identified using RepeatMasker (A. Smit and P. Green,
 unpublished.) for human and mouse sequences.
 Genes and region of sequence similarity are identified by BLAST
 (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the
 EST and cDNA sequences. Genes demonstrate at least two exons
 flanked by consensus splice sites that maintained sequence
 continuity across the splice junctions. Sequences that are not
 identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum
 standard of double strand coverage with a minimum of 2 clones and 2
 reads with no ambiguities or 2 chemistries with a minimum of 2
 clones and 3 reads with no ambiguities. If the sequence quality for
 a region does not meet this standard, it will be indicated in the
 annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality
 standards - estimated error rate less than 1 per 10,000 bases.
 Reports of lowest quality individual bases and measures of base
 quality are listed below. Description of the metrics can be found
 at URL:

<http://www.hgsc.bcm.tmc.edu:8088/quality.info/genbank.annotation.ht>
 ml.

FEATURES

Location/Qualifiers
 1. .81001
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="3"
 /clone="RP11-757F18"
 complement(1. .2041)
 /note="overlaps bases 1. .2041 of clone AC119734"
 /function="clone overlap"
 11149. .11300
 /standard_name="RH65506"
 27049. .27072
 /note="Size confirmed by PCR and restriction digest."
 /function="pcr product sequence only"
 31707. .31920
 /standard_name="G44369"
 32757. .33036
 /standard_name="D1183206"
 68551. .68832
 /standard_name="SHGC-112515"
 71827. .71979
 /standard_name="RH99175"
 74784. .75043
 /standard_name="SHGC-68798"
 77907. .78053

misc_feature

STS

misc_feature

STS

STS

STS

STS

STS

STS

/standard_name="RH93661"
77944..78073
/standard_name="SHGC-59894"
78015..81001
misc_feature
/note="overlaps bases 1..2988 of clone AC024887"
/function="clone overlap"

ORIGIN

Alignment Scores:

Pred. No.:	0.0808	Length:	81001
Score:	126.00	Matches:	30
Percent Similarity:	59.76%	Conservative:	19
Best Local Similarity:	36.59%	Mismatches:	31
Query Match:	17.62%	Indels:	2
DB:	9	Gaps:	2

US-10-005-907-2 (1-135) x AC128688 (1-81001)

QY	56	GlnGluAsnGluAsnGlySerGlySerGluGluValCysTyrThrValIleAsnHis---	74
DB	70950	CAGGACAATGTTGACCAAGACTTCTGCTGAGAGAGCTGTGCTATACCTCATCAATCATCGG	71009
QY	75	IleProHisGlnArgSerSerLeuSerSerAsnAspAspGlyTyrGluAsnIleAspSer	94
DB	71010	GTTCTCTGTACAGGCCCATCAGGGAAGTCTGCTGAGAGTACTATGAGAAATGTTCCCTGC	71069
QY	95	LeuThrArgLysValArgGlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArg	114
DB	71070	AAAGCTGAGAGACCCAGAGAGTCTTGGGAGAACTGAGACTGAGTATTCACTTCTACAT	71129
QY	115	ThrSerValSerArgProCysSerCys---ThrHisGluHisAspTyrGluValValPhe	133
DB	71130	ATGCCTTCTACAGACCCAGGCGATGCCCGATCCCGAGAGATGATATGACTTCTCATG	71189
QY	134	ProHis	135
DB	71190	CCTCAC	71195

Search completed: April 22, 2004, 17:57:43
Job time : 2245 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 22, 2004, 11:26:15 ; Search time 1784 Seconds
(without alignments)
9839.646 Million cell updates/sec

Title: US-10-005-907-1_COPY_25_429

Perfect score: 405

Sequence: 1 atgggaattatctctgcg.....atgaagtgtgtttccacac 405

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hcg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pt.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vl.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_or.*

21: em_ov.*

22: em_pat.*

23: em_ph.*

24: em_pl.*

25: em_ro.*

26: em_sts.*

27: em_un.*

28: em_vl.*

29: em_vr.*

30: em_hcg_hum.*

31: em_hcg_inv.*

32: em_hcg_other.*

33: em_hcg_mus.*

34: em_hcg_pln.*

35: em_hcg_rdt.*

36: em_hcg_mam.*

37: em_hcg_vrt.*

38: em_sy.*

39: em_hgto_hum.*

40: em_hgto_mus.*

41: em_hgto_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	405	100.0	2648	9	BC024174	BC024174 Homo sapi
2	405	100.0	3762	6	AX505122	AX505122 Sequence
3	403.4	99.6	1709	9	AK123798	AK123798 Homo sapi
4	359.4	88.7	1924	9	AB060908	AB060908 Macaca fa
5	319.2	78.8	4032	9	AK126682	AK126682 Homo sapi
6	317.6	78.4	4136	9	AK124520	AK124520 Homo sapi
7	241	59.5	141268	2	AC074365	AC074365 Homo sapi
8	241	59.5	185467	9	AL606804	AL606804 Human DNA
9	58	14.3	1532	9	BC030506	BC030506 Homo sapi
10	58	14.3	1654	9	AF521911	AF521911 Homo sapi
11	58	14.3	3270	9	AY212246	AY212246 Homo sapi
12	49	12.1	7218	6	I66494	I66494 Sequence 14
13	49	12.1	177229	2	BX855623	BX855623 Danio fer
14	49	12.1	188317	5	BX004780	BX004780 Zebrat
15	48	11.9	4218	9	AK125521	AK125521 Homo sapi
16	48	11.9	81001	9	AC128688	AC128688 Homo sapi
17	48	11.9	203405	2	AC024964	AC024964 Homo sapi
18	48	11.9	214324	2	AC110870	AC110870 Homo sapi
19	47.2	11.7	1643	6	AX740457	AX740457 Sequence
20	45.4	11.2	190882	10	AL627105	AL627105 Mouse DNA
21	44	10.9	1672	10	MMU13263	UI3263 Mus muscul
22	44	10.9	217409	2	AC130828	AC130828 Mus muscu
23	43.8	10.8	252980	10	AL669931	AL669931 Mouse DNA
24	43.4	10.7	69028	9	AC087493	AC087493 Homo sapi
25	43.4	10.7	124170	9	AC006000	AC006000 Homo sapi
26	43.4	10.7	186360	9	AC007982	AC007982 Homo sapi
27	43.4	10.7	218922	2	AC008106	AC008106 Homo sapi
28	43.2	10.7	372	4	S75319	S75319 12S rRNA [n
29	43.2	10.7	176825	10	AC132437	AC132437 Mus muscu
30	43.2	10.7	199649	10	AL669825	AL669825 Mouse DNA
31	42.8	10.6	237384	2	AC131424	AC131424 Rattus no
32	42.8	10.6	24861	2	AC106089	AC106089 Rattus no
33	42.6	10.5	221012	10	AL672234	AL672234 Mouse DNA
34	42.6	10.5	268303	2	AL772317	AL772317 Mus muscu
35	42.6	10.5	294608	2	AC129004	AC129004 Rattus no
36	42.4	10.5	151316	9	AC005039	AC005039 Homo sapi
37	42.4	10.5	173924	2	AC102756	AC102756 Mus muscu
38	42.4	10.5	193629	2	AC139376	AC139376 Mus muscu
39	42.4	10.5	212480	2	AC135862	AC135862 Mus muscu
40	42.2	10.4	185872	2	AC145959	AC145959 Gallus ga
41	42	10.4	171516	2	AC114649	AC114649 Mus muscu
42	42	10.4	195726	2	AC112925	AC112925 Mus muscu
43	41.8	10.3	500	5	DLSEARAG1	AF137202 Dicentrar
44	41.8	10.3	158510	2	AC102478	AC102478 Mus muscu
45	41.6	10.3	239	11	G37943	G37943 C4M29 Plasm

ALIGNMENTS

RESULT 1
BC024174
LOCUS BC024174
DEFINITION Homo sapiens LOC148823, mRNA (cdna clone MGC:24564 IMAGE:4109064),
complete cds.
ACCESSION BC024174
VERSION BC024174.1 GI:18848218
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2648)
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Pri 04-OCT-2003

Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Schaeetz,T.E., Brownstein,M.J., Usdin,T.B., Toehiyuki,S., Carnici,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullany,S.D., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,J.S., Krzywinski,M.I., Skalska,U., Smalish,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16999-16903 (2002)

23388257

12477932

2 (Bases 1 to 2648)

Strausberg,R.

Direct Submission

Submitted (19-FEB-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgapbs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: CLONTECH Laboratories, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)

DNA Sequencing by: Institute for Systems Biology

<http://www.systemsbio.org>

contact: amadan@systemsbiology.org

Anup Madan, Jessica Fahey, Erin Helton, Mark Kettman, Anuradha Madan, Stephanie Rodriguez, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>

Series: IMAGE Plate: 32 Row: b Column: 10

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 21687148.

FEATURES

source

1..2648

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:24564 IMAGE:4109064"

/tissue_type="Bone marrow, acute myelogenous leukemia"

/clone_lib="NIH MGC_55"

/lab_host="DH10B"

/note="Vector: pDNR-LIB"

1..2648

/gene="LOC148823"

/db_xref="LocusID:148823"

48..455

/codon_start=1

/product="LOC148823"

/protein_id="AAH24174.1"

/db_xref="GI:18848219"

/db_xref="LocusID:148823"

gene

1..2648

/gene="LOC148823"

/db_xref="LocusID:148823"

48..455

/codon_start=1

/product="LOC148823"

/protein_id="AAH24174.1"

/db_xref="GI:18848219"

/db_xref="LocusID:148823"

CDS

1..2648

/gene="LOC148823"

/db_xref="LocusID:148823"

48..455

/codon_start=1

/product="LOC148823"

/protein_id="AAH24174.1"

/db_xref="GI:18848219"

/db_xref="LocusID:148823"

/translation="MGNYLLKLSCLGENOKPKKGNPDEERKQEMTTFERKLQDQD
KKSEVSTNQENSGSEEVCTVINIHPQRSLSSNDGYNIDSLTRKVRQF
RRSETEYALLRTSVSRPCSTHEHDYEVVFPH"

ORIGIN

Query Match 100.0%; Score 405; DB 9; Length 2648;

Best Local Similarity 100.0%; Pred. No. 9.5e-94;

Matches 405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGAAATTATCTCTCGGAAATCTCAGTTGCTGGGAGAGATCAAAAGAGCCCAAG 60

DB 48 ATGGGAAATTATCTCTCGGAAATCTCAGTTGCTGGGAGAGATCAAAAGAGCCCAAG 107

QY 61 AAAGGAAATCCAGTACAGGAAAGAAACCGGAGGAATGCTACATTTGAAAGAAACTT 120

DB 108 AAAGGAAATCCAGTACAGGAAAGAAACCGGAGGAATGCTACATTTGAAAGAAACTT 167

QY 121 CAAGATCAAGATAAGAAAGCAAGAACTTTCATCCCTTCTAATCAGGAAACGAGAAT 180

DB 168 CAAGATCAAGATAAGAAAGCAAGAACTTTCATCCCTTCTAATCAGGAAACGAGAAT 227

QY 181 GCAGTGGTTCTGAAGAAAGTGTCTACTGTCTATTAATCATCATCCCCCATCAGATCC 240

DB 228 GCAGTGGTTCTGAAGAAAGTGTCTACTGTCTATTAATCATCATCCCCCATCAGATCC 287

QY 241 TCCTTGAGTCCCAATGATGATGGCTATCAGAAACATTTGACTCCCTCACAAGGAAAGTGA 300

DB 288 TCCTTGAGTCCCAATGATGATGGCTATCAGAAACATTTGACTCCCTCACAAGGAAAGTGA 347

QY 301 CAGTTTAGGAAAGTTCAGACAGAAATATGCCCTTCTTAGGACTTCTGTAGTAGGCCT 360

DB 348 CAGTTTAGGAAAGTTCAGACAGAAATATGCCCTTCTTAGGACTTCTGTAGTAGGCCT 407

QY 361 TGTTCCTCGCACCCTCAGCATGATTTATCAAGTTGTGTTTCCACAC 405

DB 408 TGTTCCTCGCACCCTCAGCATGATTTATCAAGTTGTGTTTCCACAC 452

RESULT 2

AX505122

LOCUS AX505122 3762 bp DNA linear PAT 27-SEP-2002

DEFINITION Sequence 1 from Patent WO0246389.

ACCESSION AX505122

VERSION AX505122.1 GI:23386429

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1

NOCKA,K., PIROZZI,G. and EINSTEIN,R.

Novel genes associated with allergic hypersensitivity and mast cell activation

Patent: WO 0246389-A 13-JUN-2002;

UCB, S.A. (BE)

FEATURES

Location/Qualifiers

1..3762

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

25..432

/note="unnamed protein product"

/codon_start=1

/protein_id="CAD48808.1"

/db_xref="GI:23386430"

/db_xref="REMBL:CAD48808"

/translation="MGNYLLKLSCLGENOKPKKGNPDEERKQEMTTFERKLQDQD
KKSEVSTNQENSGSEEVCTVINIHPQRSLSSNDGYNIDSLTRKVRQF
RRSETEYALLRTSVSRPCSTHEHDYEVVFPH"

ORIGIN

Query Match 100.0%; Score 405; DB 6; Length 3762;

Best Local Similarity 100.0%; Pred. No. 9.1e-94;

Matches 405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGAAATTATCTCTCGGAAATCTCAGTTGCTGGGAGAGATCAAAAGAGCCCAAG 60

DB 25 ATGGGAAATTATCTCTCGGAAATCTCAGTTGCTGGGAGAGATCAAAAGAGCCCAAG 84

QY 61 AAAGGAAATCCAGTACAGGAAAGAAACCGGAGGAATGCTACATTTGAAAGAAACTT 120

DB 85 AAAGGAAATCCAGTACAGGAAAGAAACCGGAGGAATGCTACATTTGAAAGAAACTT 144

QY 121 CAAGATCAAGTAAGAAAGCCAGAGTTTTCATCCACTTCTTAATCAGGAAACGAGAAAT 180
 Db 145 CAAGATCAAGTAAGAAAGCCAGAGTTTTCATCCACTTCTTAATCAGGAAACGAGAAAT 204
 QY 181 GCGAGTGGTCTCTGAAGAGTGTGTACACTGTGTCAATTAATCAGATCCATCCCATCAGAGATCC 240
 Db 205 GCGAGTGGTCTCTGAAGAGTGTGTACACTGTGTCAATTAATCAGATCCATCCCATCAGAGATCC 264
 QY 241 TCCCTGAGCTCCCAATGATGGCTATGAGAACATTCGACTCCCTCAAGAGAAAGTGAGA 300
 Db 265 TCCCTGAGCTCCCAATGATGGCTATGAGAACATTCGACTCCCTCAAGAGAAAGTGAGA 324
 QY 301 CAGTTTAGAGAAAGGTTCAGACAGAAATATGCCCTTCTTAGGACTTCTGTAGTAGGCTT 360
 Db 325 CAGTTTAGAGAAAGGTTCAGACAGAAATATGCCCTTCTTAGGACTTCTGTAGTAGGCTT 384
 QY 361 TGTTCCTGCACCCATGAGCATGATTATGAAGTGTGTTCACAC 405
 Db 385 TGTTCCTGCACCCATGAGCATGATTATGAAGTGTGTTCACAC 429
 RESULT 3
 AK123798 1709 bp mRNA linear PRI 09-SEP-2003
 LOCUS Homo sapiens cDNA FLJ1804 fis, clone NOVAR2000710.
 DEFINITION
 VERSION AK123798.1 GI:34529425
 KEYWORDS oligo capping; fis (full insert sequence).
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1
 Suzuki, O., Sasaki, N., Aotseuka, S., Shoji, T., Ichihara, T.,
 Shiohata, N., Matsumoto, K., Hirano, M., Sano, S., Nomura, R.,
 Yoshikawa, Y., Matsumura, Y., Moriya, S., Chiba, E., Momiyama, H.,
 Onogawa, S., Kaeriyama, S., Sato, N., Matsunawa, H., Takahashi, E.,
 Kataoka, R., Kuga, N., Kuroda, A., Sato, I., Kanata, K., Takami, S.,
 Tarashima, Y., Watanabe, M., Sugiyama, T., Irie, R., Otsuki, T.,
 Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y.,
 Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H.,
 Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K.,
 Wagatsuma, M., Murakawa, K., Kanehori, K., Takahashi-Fujii, A.,
 Ohshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S.,
 Nagahara, K., Masuho, Y., Nagai, K. and Isogai, T.
 NEDO human cDNA sequencing project
 Unpublished
 2 (bases 1 to 1709)
 Isogai, T. and Yamamoto, J.
 Direct Submission
 Submitted (15-JUL-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7
 Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan
 (E-mail: genonice@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
 NEDO human cDNA sequencing project supported by Ministry of
 Economy, Trade and Industry of Japan; cDNA full insert sequencing:
 Research Association for Biotechnology (RAB); cDNA library
 construction: Helix Research Institute (HRI) (supported by Japan
 Key Technology Center etc.); 5' - & 3' - end one pass sequencing: RAB,
 HRI, and Biotechnology Center, National Institute of Technology and
 Evaluation; clone selection for full insert sequencing: HRI and
 RAB; annotation: HRI and RAB.
 Location/Qualifiers
 1. .1709
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="NOVAR2000710"
 /tissue_type="ovary"
 /clone_lib="NOVAR2"
 /dev_stage="adult"
 /note="cloning vector: pME18SFL3"
 FEATURES
 source
 ORIGIN

Query Match 99.6%; Score 403.4; DB 9; Length 1709;
 Best Local Similarity 99.8%; Pred. No. 2.6e-93;
 Matches 404; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 ATGGGAAATATATCTCTCGGAAATCTCAGTTGCTGGGAGAGAAATCAAAAAGGCCCAAG 60
 Db 59 ATGGGAAATATATCTCTCGGAAATCTCAGTTGCTGGGAGAGAAATCAAAAAGGCCCAAG 118
 QY 61 AAAGGAAACCCAGATGAGGAAAGAAACGGCAGGAAATGACTACATTTTGAAGAAAACCTT 120
 Db 119 AAAGGAAACCCAGATGAGGAAAGAAACGGCAGGAAATGACTACATTTTGAAGAAAACCTT 178
 QY 121 CAAGATCAAGTAAGAAAGCCAGAAAGTTTCATCCACTTCTTAATCAGGAAACGAGAAAT 180
 Db 179 CAAGATCAAGTAAGAAAGCCAGAAAGTTTCATCCACTTCTTAATCAGGAAACGAGAAAT 238
 QY 181 GCGAGTGGTCTCTGAAGAGTGTGTACACTGTCTTAATCAGATCCATCCCATCAGAGATCC 240
 Db 239 GCGAGTGGTCTCTGAAGAGTGTGTACACTGTCTTAATCAGATCCATCCCATCAGAGATCC 298
 QY 241 TCCCTGAGCTCCCAATGATGGCTATGAGAACATTCGACTCCCTCAAGGAAAGTGAGA 300
 Db 299 TCCCTGAGCTCCCAATGATGGCTATGAGAACATTCGACTCCCTCAAGGAAAGTGAGA 358
 QY 301 CAGTTTAGAGAAAGGTTCAGACAGAAATATGCCCTTCTTAGGACTTCTGTAGTAGGCTT 360
 Db 359 CAGTTTAGAGAAAGGTTCAGACAGAAATATGCCCTTCTTAGGACTTCTGTAGTAGGCTT 418
 QY 361 TGTTCCTGCACCCATGAGCATGATTATGAAGTGTGTTCACAC 405
 Db 419 TGTTCCTGCACCCATGAGCATGATTATGAAGTGTGTTCACAC 463
 RESULT 4
 AB060908 1924 bp mRNA linear PRI 10-JAN-2002
 LOCUS Macaca fascicularis brain cDNA clone: Qtra-14007, full insert
 DEFINITION sequence.
 ACCESSION AB060908
 VERSION AB060908.1 GI:13874585
 KEYWORDS oligo capping; fis (full insert sequence).
 SOURCE Macaca fascicularis (crab-eating macaque)
 ORGANISM Macaca fascicularis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
 Cercopitheinae; Macaca.
 1
 Osada, N., Hida, M., Kusuda, J., Tanuma, R., Hirata, M., Hirai, M.,
 Terao, K., Suzuki, Y., Sugano, S. and Hashimoto, K.
 Prediction of unidentified human genes on the basis of sequence
 similarity to novel cDNAs from cynomolgus monkey brain
 Online Publication
 Genome Biology 2001 3 (1): research0006.1-0006.5;
 http://genomebiology.com/2001/3/1/research/0006/
 2 (bases 1 to 1924)
 Hashimoto, K., Osada, N., Hida, M., Kusuda, J. and Sugano, S.
 Direct Submission
 Submitted (27-APR-2001) Katsuyuki Hashimoto, National Institute of
 Infectious Diseases, Division of Genetic Resources; 23-1, Toyama
 1-chome, Shijuku-ku, Tokyo 162-8640, Japan
 (E-mail: khashi@nih.go.jp, URL: http://www.nih.go.jp/yoken/genebank/,
 Tel: 81-3-5285-1111 (ex. 2120), Fax: 81-3-5285-1181)
 Lab host: TOP10
 Vector: pME18S-FL3 (Acc. No. AB009864)
 R. Site1: DraIII (CAGCATGTG)
 R. Site2: DraIII (CAGCATGTG)
 Description: 1st strand cDNA was primed with an oligo(dT) primer
 [ATGCGCCCTTTTCTTTTCTTTT]; double-stranded cDNA was synthesized
 using specific 5' and 3' primers and amplified by PCR. The PCR
 product was digested with SfiI and size selection was performed to
 exclude fragments <1.5kb. The SfiI-digested PCR product was cloned
 into distinct DraIII sites of pME18S-FL3. XhoI sites just outside

the DraIII sites can be used to isolate the cDNA insert. Libraries were constructed by oligo-capping method (Sugano et al., Institute of Medical Science, University of Tokyo).

Custom primer used for sequencing
(5' end primer [CTCTGCTCAAGACTGGG];
3' end primer [CGACCTGCTCAGGACCA]).

FEATURES

source

1. 1924

/organism="Macaca fascicularis"

/mol_type="mRNA"

/db_xref="taxon:9541"

/clone="Qcra-14007"

/sex="male"

/tissue_types="temporal lobe right"

/clone_lib="macaque brain cDNA library Qcra"

/dev_stage="adult"

248..559

/codon_start=1

/product="hypothetical protein"

/protein_id="BA846903.1"

/db_xref="GI:13874586"

/translations="MTTFERKLQDDKSKVESSISQENENGSGSEEVCTVIVNHIP

HRSSLSNDGVENIDSLTRKREFRSETEYALLRTSVSRPCTTHEHDYEVVLP

H"

ORIGIN

Query Match 88.7%; Score 359.4; DB 9; Length 1924;
Best Local Similarity 94.7%; Pred. No. 5.5e-82;
Matches 372; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 13 CTCCTCGGMAAATCAGTTGCTGGGAGAGAAATCAAAAGAACGCCAAGAAAGAAACCCCA 72
DB 164 CGCGGAGGACATGAGTTGCTTGGGAGAGAAATCAAAAGAACGCCAAGAAAGAAACCCCA 223
QY 73 GATGAGGAAGAAACCGGAGGAAATGACTATCAATTTGAAGAAATCTCAAGATCAAGAT 132
DB 224 GATGAGGAAGAAACCGGAGGAAATGACTATCAATTTGAAGAAATCTCAAGATCAAGAT 283
QY 133 AAGAAAGCCCAAGAAAGTTTCAATCAGGAAACGAGAAATGAGTGGTGTCT 192
DB 284 AAGAAAGCAAGAAAGTTTCAATCAGGAAACGAGAAATGAGTGGTGTCT 343
QY 193 GAAGAAAGTGCTACATGCTATTAATCAATCAATCCCATCAGAGATCCCTCCCTGAGCTCC 252
DB 344 GAAGAAAGTGCTACATGCTATTAATCAATCAATCCCATCAGAGATCCCTCCCTGAGCTCC 403
QY 253 AATGATGCTGATGAGAAATGACTCCCTCAGAGAAATGAGTGGTGTCT 312
DB 404 AATGATGCTGATGAGAAATGACTCCCTCAGAGAAATGAGTGGTGTCT 463
QY 313 AGGTGAGAGACAGAAATGCTTTTGAAGATCTGTTAGAGCTTTGTTAGGCTTTGTTCCGACC 372
DB 464 AGGTGAGAGACAGAAATGCTTTTGAAGATCTGTTAGAGCTTTGTTAGGCTTTGTTCCGACC 523
QY 373 CATGAGCATGATGAGTGGTGTGTTCCACAC 405
DB 524 CATGAGCATGATGAGTGGTGTGTTCCACAC 556

RESULT 5

AKI26682

LOCUS

AKI26682 Homo sapiens cDNA FLJ44728 fis, clone BRACE3024537. 4032 bp mRNA linear PRI 09-SEP-2003

DEFINITION

AKI26682

ACCESSION

AKI26682.1

VERSION

AKI26682.1

KEYWORDS

oligo capping; fis (full insert sequence).

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1

Authors

Kodaira, H., Furuya, T., Takahashi, M., Kikkawa, E., Omura, Y., Abe, K.,
Kanihara, K., Katsuka, N., Sato, K., Tanikawa, M., Yamazaki, M.,
Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S.,
Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T.,
Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M.,
Kikuchi, H., Murakawa, K., Kanenori, K., Takahashi-Fujii, A.,
Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S.,
Nagahara, K., Maeuho, Y., Nagai, K. and Isogai, T.
NEO human cDNA sequencing project

Unpublished

2 (bases 1 to 4032)

Isogai, T. and Yamamoto, J.

Direct Submission

TITLE

JOURNAL

Submitted (15-JUL-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7

Karusa-Kamatari, Kisarazu, Chiba 292-0818, Japan

(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)

NEO human cDNA sequencing project supported by Ministry of

Economy, Trade and Industry of Japan; cDNA full insert sequencing:

Research Association for Biotechnology (RAB); cDNA library

construction: Helix Research Institute (HRI) (supported by Japan

Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,

HRI, and Biotechnology Center, National Institute of Technology and

Evaluation; clone selection for full insert sequencing: HRI and

RAB; annotation: HRI and RAB.

FEATURES

Location/Qualifiers

1. 4032

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="BRACE3024537"

/tissue_type="cerebellum"

/clone_lib="BRACE3"

/note="Cloning vector: pME18SFL3"

ORIGIN

Query Match 78.8%; Score 319.2; DB 9; Length 4032;
Best Local Similarity 94.8%; Pred. No. 1.1e-71;
Matches 330; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 58 AAGAAAGAAACCCAGATGAGGAAAGAAACGCGAGGAAATGACTATTTGAAGAAAA 117
DB 374 AAGTGGGAATTTATCTCTCGGAATCTCAGGAGGAAATGACTATTTGAAGAAAA 433
QY 118 CTTCAAGATCAAGTAAAGAAAGCCAAAGAGTTTCATCCACTTCTAATTCAGAAAAAGAG 177
DB 434 CTTCAAGATCAAGTAAAGAAAGCCAAAGAGTTTCATCCACTTCTAATTCAGAAAAAGAG 493
QY 178 AATGCACTGCTTCTGAGAGAGTGTCTACACTCTCATTAATCATCCCCCATCAGAGA 237
DB 494 AATGCACTGCTTCTGAGAGAGTGTCTACACTCTCATTAATCATCCCCCATCAGAGA 553
QY 238 TCCTCCCTGAGCTCCAATGATGATGGCTATGAGAACTTGACTCCCTCACAAGAAAAAGTG 297
DB 554 TCCTCCCTGAGCTCCAATGATGATGGCTATGAGAACTTGACTCCCTCACAAGAAAAAGTG 613
QY 298 AGACAGTTTAGAGAAAGGTGAGAGACAGCAATATGCCCTTCTTAGAGCTTCTGTTAGTAGG 357
DB 614 AGACAGTTTAGAGAAAGGTGAGAGACAGCAATATGCCCTTCTTAGAGCTTCTGTTAGTAGG 673
QY 358 CTTGTTCTTCGCCACCCATGAGCATGATTATGAAGTTGTGTTTCCACAC 405
DB 674 CTTGTTCTTCGCCACCCATGAGCATGATTATGAAGTTGTGTTTCCACAC 721

RESULT 6

AKI24520

LOCUS

AKI24520

DEFINITION

AKI24520

ACCESSION

AKI24520.1

VERSION

AKI24520.1

KEYWORDS

oligo capping; fis (full insert sequence).

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

AKI24520

Homo sapiens cDNA FLJ42529 fis, clone BRACE3002390. 4136 bp mRNA linear PRI 09-SEP-2003

DEFINITION

AKI24520

ACCESSION

AKI24520.1

VERSION

AKI24520.1

KEYWORDS

oligo capping; fis (full insert sequence).

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Ninomiya, K., Wagatsuma, M., Kanda, K., Kondo, H., Yokoi, T.,
Kodaira, H., Furuuya, T., Takahashi, M., Kikkawa, B., Omura, Y., Abe, K.,
Kanihara, K., Katsuta, N., Sato, K., Tanikawa, M., Yanazaki, M.,
Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S.,
Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T.,
Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M.,
Kikuchi, H., Murakawa, K., Kaneshiro, K., Takahashi-Fujii, A.,
Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S.,
Nagahara, K., Masuko, Y., Nagai, K. and Isogai, T.
NEO human cDNA sequencing project
Unpublished
2 (bases 1 to 4136)
Isogai, T. and Yamamoto, J.
Direct Submission
Submitted (15-JUL-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7
Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
NEO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5' - & 3' - end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.
FEATURES
Location/Qualifiers
1..4136
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="BRACE3002390"
/tissue_type="cerebellum"
/clone_lib="BRACE3"
/note="cloning vector: pME18SFL3"
136..537
/note="unnamed protein product"
/codon_start=1
/protein_id="BAC85873.1"
/db_xref="GI:34530325"
/translation="MWMLEDSNSTGCGARNCVCLVFLSVLGCQSKKQMTQQAG
RWLRAGEASSETPEGLDFOSELPFLAAGWQHPVRCNVLPHAFSSWCFQG
NFPKWSQELNRVTVKRWELIISGSRK"
ORIGIN
Query Match 78.4%; Score 317.6; DB 9; Length 4136;
Best Local Similarity 94.5%; Pred. No. 2.9e-71;
Matches 329; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
QY 58 AAGAAAGGAAACCCAGATGAGGAAAGAAACGGCAGGAAATGACTATCATTTTGAAGAAA 117
DB 495 AAGATGGGAATATTCTCTCTCGAAACTCAGGCGAGAAATGACTATCATTTTGAAGAAA 554
QY 118 CTTCAAGATCAAGTAAGAAAGCCAGAGTTTCATCCATTTCTATCAGGAAACGAG 177
DB 555 CTTCAAGATCAAGTAAGAAAGCCAGAGTTTCATCCATTTCTATCAGGAAACGAG 614
QY 178 AATGCGACTGGTCTTGAAGAAAGTGTGCTACACTGTTCATTAATCATATCATCCCATCAGAGA 237
DB 615 AATGCGACTGGTCTTGAAGAAAGTGTGCTACACTGTTCATTAATCATATCATCCCATCAGAAA 674
QY 238 TCCTCCCTGAGTCCAGTATGATGGCTTATGAGAACATTCAGTCCCTCAGAGAAAGTG 297
DB 675 TCCTCCCTGAGTCCAGTATGATGGCTTATGAGAACATTCAGTCCCTCAGAGAAAGTG 734
QY 298 AGACAGTTTAGAAGAGTTCAGAGACAGAAATATGCCCTTCTTAGCACTTCTGTAGTAGG 357
DB 735 AGACAGTTTAGAAGAGTTCAGAGACAGAAATATGCCCTTCTTAGCACTTCTGTAGTAGG 794
QY 358 CTTGTTCTTCCGACCCATGAGCATGATTATGAAGTTGTGTTCCACAC 405
|||||

Db
RESULT 7
AC074365/c
LOCUS
DEFINITION
AC074365
AC074365
HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

795 CCTGTCTCTGCACCATGAGCATGATTATGAAGTTGTGTTCCACAC 842
AC074365
Homo sapiens chromosome 1 clone RP11-115C4, WORKING DRAFT SEQUENCE,
10 unordered pieces.
AC074365
HTG; HTGS_PHASE1; HTGS_DRAFT.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Waterston, R.H.
The sequence of Homo sapiens clone
Unpublished
2 (bases 1 to 141268)
Waterston, R.H.
Direct Submission
Submitted (29-JUL-2000) Genome Sequencing Center, Washington
University School of Medicine, 444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Sep 23, 2000 this sequence version replaced gi:9838075.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H.NH0115C04
----- Summary Statistics -----
Sequencing vector: pMI3; 100%
Sequencing strategy: Dye-terminator Big Dye; 0% of reads
Chemistry: Dye-terminator Big Dye; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 136274 bases at least Q40
Consensus quality: 138356 bases at least Q30
Consensus quality: 139182 bases at least Q20
Insert size: 142000; agarose-fp
Insert size: 140368; sum-of-contigs
Quality coverage: 4.64 in Q20 bases; agarose-fp
Quality coverage: 4.81 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 10 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 31377: contig of 31377 bp in length
* 31378 31477: gap of unknown length
* 31478 63011: contig of 31534 bp in length
* 63011 63112: gap of unknown length
* 63112 64845: contig of 1734 bp in length
* 64845 64945: gap of unknown length
* 64945 67606: contig of 2661 bp in length
* 67606 67707: gap of unknown length
* 67707 77674: contig of 9968 bp in length
* 77674 89256: contig of 11482 bp in length
* 89256 89357: gap of unknown length
* 89357 100400: contig of 11044 bp in length
* 100400 100501: gap of unknown length
* 100501 111915: contig of 11415 bp in length
* 111915 125713: contig of 13698 bp in length
* 125713 125813: gap of unknown length
* 125813 125914: gap of unknown length

* 125814 141268: contig of 15455 bp in length.

```

FEATURES
  source
    Location/Qualifiers
      1..141268
        /organism="Homo sapiens"
        /mol_type="genomic DNA"
        /db_xref="taxon:9606"
        /chromosome="1"
        /clone="RP11-115C4"
      misc_feature
        1..31377
        /note="assembly_name:Contig10"
      misc_feature
        31478..63011
        /note="assembly_name:Contig11"
      misc_feature
        63112..64845
        /note="assembly_name:Contig2"
      misc_feature
        64946..67606
        /note="assembly_name:Contig3"
      misc_feature
        67707..77674
        /note="assembly_name:Contig4"
      misc_feature
        77775..89256
        /note="assembly_name:Contig5"
      misc_feature
        89357..100400
        /note="assembly_name:Contig6"
      misc_feature
        100501..111915
        /note="assembly_name:Contig7"
      misc_feature
        112016..125713
        /note="assembly_name:Contig8"
      misc_feature
        125814..141268
        /note="assembly_name:Contig9"
        clone_end:SP6
        vector_side:left"
  ORIGIN
    Query Match      59.5%; Score 241; DB 2; Length 141268;
    Best Local Similarity 96.1%; Pred. No. 9.4e-52;
    Matches 247; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
  QY 149 TTTCATCCACTTCTTAATCAGGAAACGAGATGGCAGTGGTTCCTGAAGAAGTGTGCTACA 208
  Db 5717 TTTCCTTTGTGCTTGGCAGGAAACGAGATGGCAGTGGTTCCTGAAGAAGTGTGCTACA 5658
  QY 209 CTGTCATTAATCACAATCCCAATCAGAGATCTCTCCAGTCCCAATGATGATGCTATG 269
  Db 5657 CTGTCATTAATCACAATCCCAATCAGAGATCTCTCCAGTCCCAATGATGATGCTATG 5598
  QY 269 AGAATCATTGACCTCCCTCAAGGAAAGTGAGACATCTTTAGAGAAAGGTGACAGACAGAT 328
  Db 5597 AGAATCATTGACCTCCCTCAAGGAAAGTGAGACATCTTTAGAGAAAGGTGACAGACAGAT 5538
  QY 329 ATGCCCTTCTTAGGACTTCTGTAGTAGGCCTTGTCTCCAGTCCCAATGATGATGCTATG 388
  Db 5537 ATGCCCTTCTTAGGACTTCTGTAGTAGGCCTTGTCTCCAGTCCCAATGATGATGCTATG 5478
  QY 389 AAGTTGTGTTTCCACAC 405
  Db 5477 AAGTTGTGTTTCCACAC 5461

```

```

RESULT 8
AL606804
LOCUS
DEFINITION Human DNA sequence from clone RP11-978115 on chromosome 1, complete
sequence.
ACCESSION AL606804
VERSION AL606804.11 GI:24366459
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 185467)
AUTHORS Almeida, J.
TITLE Direct Submission
JOURNAL Submitted (16-OCT-2002) Wellcome Trust Sanger Institute, Hinxton,

```

Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
 On Oct 24, 2002 this sequence version replaced gi:18121511.
 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at

http://www.sanger.ac.uk/HGP/Chr1
 RP11-978115 is from the library RPCI-11.4 constructed by the group of Pieter de Jong. For further details see
 http://www.chori.org/bacpac/home.htm
 VECTOR: pBAC3.6

----- Genome Center
 Center: Wellcome Trust Sanger Institute
 Center code: SC
 Web site: http://www.sanger.ac.uk
 Contact: humquery@sanger.ac.uk

FEATURES

Location/Qualifiers
 source
 1..185467
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="1"
 /clone="RP11-978115"
 /clone_lib="RPCI-11.4"

ORIGIN

```

Query Match      59.5%; Score 241; DB 9; Length 185467;
Best Local Similarity 96.1%; Pred. No. 9e-52;
Matches 247; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
  QY 149 TTTCATCCACTTCTTAATCAGGAAACGAGATGGCAGTGGTTCCTGAAGAAGTGTGCTACA 208
  Db 132063 TTTCCTTTGTGCTTGGCAGGAAACGAGATGGCAGTGGTTCCTGAAGAAGTGTGCTACA 132122
  QY 209 CTGTCATTAATCACAATCCCAATCAGAGATCTCTCCAGTCCCAATGATGATGCTATG 268
  Db 132123 CTGTCATTAATCACAATCCCAATCAGAGATCTCTCCAGTCCCAATGATGATGCTATG 132182
  QY 269 AGAATCATTGACCTCCCTCAAGGAAAGTGAGACATCTTTAGAGAAAGGTGACAGACAGAT 328
  Db 132183 AGAATCATTGACCTCCCTCAAGGAAAGTGAGACATCTTTAGAGAAAGGTGACAGACAGAT 132242
  QY 329 ATGCCCTTCTTAGGACTTCTGTAGTAGGCCTTGTCTCCAGTCCCAATGATGATGCTATG 388
  Db 132243 ATGCCCTTCTTAGGACTTCTGTAGTAGGCCTTGTCTCCAGTCCCAATGATGATGCTATG 132302
  QY 389 AAGTTGTGTTTCCACAC 405
  Db 132303 AAGTTGTGTTTCCACAC 132319

```

RESULT 9
 BC030506
 LOCUS

BC030506 1532 bp mRNA linear PRI 06-OCT-2003

DEFINITION Homo sapiens germinal center expressed transcript 2, mRNA (CDNA clone MGC:40441 IMAGE:4385178), complete cds.

ACCESSION BC030506

VERSION BC030506.1 GI:20987805

KEYWORDS MGC.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1532) Strausberg, R.L., Feigold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Schmen, C.W., Schuler, G.D., Altschul, S.F., Zebberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, P., Diatchenko, L., Marudina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.P., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Uddin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McSwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S.D., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Rahe, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smal, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

TITLE

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

MEDLINE 2238257

PUBMED 12477932

REFERENCE 2 (bases 1 to 1532) Strausberg, R.

AUTHORS Direct Submission

TITLE Submitted (07-MAY-2002) National Institutes of Health, Mammalian Gene Collection (MGC) Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT Contact: MGC help desk Email: cgabs-remail.nih.gov Tissue Procurement: Lou Staud cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland Web site: <http://www.nisc.nih.gov/> Contact: nisc_mgc@hgrl.nih.gov Akhter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Heghigni, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.L., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stanterop, S., Thomas, P.J., Touchman, J.W., Taurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov> Series: TRAK Plate: 64 Row: a Column: 15 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 22749536. Location/Qualifiers

1..1532

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:40441 IMAGE:4385178"

/tissue_type="Lymph, lymphoma"

/clone_lib="NTH_MGC_85"

gene

1..1532

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

/gene="GCET2"

/note="synonyms: HGAL, GCAT2, MGC40441"

/db_xref="locusid:257144"

/db_xref="MIM:607792"

54..590

/codon_start=1

/product="germinal center expressed transcript 2"

/protein_id="AAH30506.1"

/db_xref="GI:20987805"

/db_xref="locusid:257144"

/translation="MGSLRLNRNRQNTQEMFWNQSPKORTSRCHWHTAEGCF CLPKKILIFEKQDSQNERMSTPIQNDQTYSEELCYTLINHRVICTIPSGMS AEEYENVPCKAERPRESLGTTETYSLLHMPSTDPHAKSPEDYELLMPHRLISHP LQPRPLMAPSETQFSLH"

CDS

Query Match 14.3%; Score 58; DB 9; Length 1532; Best Local Similarity 54.4%; Pred. No. 0.00023; Matches 162; Conservative 0; Mismatches 130; Indels 6; Gaps 2;

QY 114 AAAACTTCAAGATCAAGATAGAAAGCAAGAGTTTCATCCACTTCTTAATCAGGAAA 173

Db 218 AAAGAGGCAAGATTCCTCCAAACGAAATGAAGATGTCACTACTCCATCCAGGACAA 277

QY 174 CGAGATGGCAGTGGTCTTGAAGAAGTGTCTACACTGTCTAATATCACA---TCCCCCA 230

Db 278 TGTTCACAGACCTACTCAGAGGAGCTGTGTATACCTCATCAATCCTGGTCTCTG 337

QY 231 TCAGAGATCTCCCTGAGCTCCAAATGATGATGGTGTATGAGAACATTCAGTCCCTCACAAG 290

Db 338 TACAAGGCATCAGGGAACCTCTGCTGAAGACTACTATGAGAATGTTCCCTGCAAGCTGA 397

QY 291 GAAAGTGAAGACAGTTTGAAGAAAGTTCAGACAGATATGCCCTTCT---TAGACTTC 347

Db 398 GAGACCCAGAGAGTCTCTGGGAGGAACCTGAGACTGAGTATTCCTTACATATGCTTC 457

QY 348 TGTTAGTAGGCTTCTCTCTGCACCATGAGCATGATGATGAAGTTGTGTTTCCACAC 405

Db 458 TACAGACCCAGGCAATGCCCATCCCCAGAGATGATGATGATCTCTCATGCTCTCAC 515

RESULT 10

AF521911

LOCUS 1654 bp mRNA linear PRI 14-JAN-2003

DEFINITION Homo sapiens HGAL mRNA, complete cds.

ACCESSION AF521911

VERSION AF521911.1 GI:27733682

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1654)

Lossoos, I.S., Allizadeh, A., Rajapaksa, R., Tibshirani, R. and Levy, R. HGAL is a novel interleukin-4-inducible gene that strongly predicts survival in diffuse large B-cell lymphoma

JOURNAL Blood 101 (2), 433-440 (2003)

MEDLINE 22397484

PUBMED 12509382

REFERENCE 2 (bases 1 to 1654)

Lossoos, I.S., Allizadeh, A.A. and Levy, R.

AUTHORS Direct Submission

TITLE Submitted (16-JUN-2002) Division of Oncology, Department of Medicine, Stanford University, 269 Campus Drive, CCSR building, Room 1100, Palo Alto, CA 94305-5151, USA

JOURNAL Location/Qualifiers

1..1654

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

FEATURE	source	Location/Qualifiers
*	44798	44897: gap of 100 bp
*	44898	56004: contig of 11107 bp in length
*	56005	56104: gap of 100 bp
*	56105	73586: contig of 17482 bp in length
*	73587	73686: gap of 100 bp
*	73687	76188: contig of 2502 bp in length
*	76189	76288: gap of 100 bp
*	76289	82092: contig of 5804 bp in length
*	82093	82192: gap of 100 bp
*	82193	97210: contig of 15018 bp in length
*	97211	97310: gap of 100 bp
*	97311	103462: contig of 6152 bp in length
*	103463	103562: gap of 100 bp
*	103563	107200: contig of 3638 bp in length
*	107201	107300: gap of 100 bp
*	107301	111003: contig of 3703 bp in length
*	111004	111103: gap of 100 bp
*	111104	128252: contig of 17149 bp in length
*	128253	128352: gap of 100 bp
*	128353	134350: contig of 5998 bp in length
*	134351	134450: gap of 100 bp
*	134451	141346: contig of 6896 bp in length
*	141347	141446: gap of 100 bp
*	141447	146748: contig of 5302 bp in length
*	146749	146848: gap of 100 bp
*	146849	149324: contig of 2476 bp in length
*	149325	149424: gap of 100 bp
*	149425	152942: contig of 3518 bp in length
*	152943	153042: gap of 100 bp
*	153043	164523: contig of 11487 bp in length
*	164530	164629: gap of 100 bp
*	164630	172404: contig of 7775 bp in length
*	172405	172504: gap of 100 bp
*	172505	174733: contig of 2229 bp in length
*	174734	174833: gap of 100 bp
*	174834	177229: contig of 2396 bp in length
FEATURES		
source		
1..177229		
/organism="Danio rerio"		
/mol_type="genomic DNA"		
/db_xref="taxon:7955"		
/clone="DKEX-283P18"		
/clone_lib="DanioKey"		
misc_feature	1..3652	/note="assembly fragment:00274 fragment chain:1 clone end:SP6 vector side:left"
misc_feature	3753..18203	/note="assembly fragment:00680 fragment chain:1"
misc_feature	18304..31101	/note="assembly fragment:00799 fragment chain:1"
misc_feature	31202..34797	/note="assembly fragment:00186 fragment chain:2"
misc_feature	34898..44797	/note="assembly fragment:00455 fragment chain:2"
misc_feature	44898..56004	/note="assembly fragment:00604 fragment chain:2"
misc_feature	56105..73586	/note="assembly fragment:01050 fragment chain:2"
misc_feature	73687..76188	/note="assembly fragment:00067 fragment chain:2"
misc_feature	76289..82092	/note="assembly fragment:00162 fragment chain:2"
misc feature	82193..97210	

misc_feature /note="assembly fragment:00920
fragment chain:2"
97311..103462
misc_feature /note="assembly fragment:00210
fragment chain:2"
103563..107200
misc_feature /note="assembly fragment:00054
fragment chain:2"
107301..111003
misc_feature /note="assembly fragment:00117
fragment chain:2"
11104..128252
misc_feature /note="assembly fragment:01194
fragment chain:2"
128353..134350
misc_feature /note="assembly fragment:00307
fragment chain:2"
134351..141346
misc_feature /note="assembly fragment:00340
fragment chain:2"
14147..146748
misc_feature /note="assembly fragment:00242
fragment chain:2"
146849..149324
misc_feature /note="assembly fragment:00023
fragment chain:2"
149425..152942
misc_feature /note="assembly fragment:00098
fragment chain:2"
153043..164529
misc_feature /note="assembly fragment:00529
fragment chain:2"
164630..172404
misc_feature /note="assembly fragment:00392
fragment chain:3"
172505..174733
misc_feature /note="assembly fragment:00082
fragment chain:3"
174834..177229
misc_feature /note="assembly fragment:00044"

ORIGIN

Query Match 12.1%; Score 49; DB 2; Length 177229;

Best Local Similarity 59.9%; Pred. No. 0.025;

Matches 82; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 1 ATGGGAATTTATCTCTCGGAAATCTAGTTGCCTGGGAGAGATCAAAAGAGCCCAAG 60
DB 144111 AAGACAATAACCACTCGATACACTGAGTTGTGAGAGAGCATCTCTAAAGCACAC 144170
QY 61 AAAGGAACCCAGATCAGAGAAAGAAACGGCAGGAAATGACTATATTGAAAGAAACTT 120
DB 144171 ACATGAACCTTGAGCGGATGGCTACAGCGCAAGACCACTCTGTGGCTAAGAA 144230
QY 121 CAAGATCAAGATAAGAA 137
DB 144231 CAGGAACAGGCAAGAA 144247

RESULT 14

EX004780/c BX004780 188317 bp DNA linear VRT 13-MAR-2003
LOCUS Zebrafish DNA sequence from clone CH211-202P1 in linkage group 20,
DEFINITION complete sequence.
ACCESSION BX004780
VERSION BX004780.4 GI:28552191
KEYWORDS HTG
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
REFERENCE 1 (bases 1 to 188317)

AUTHORS

Direct Submission
Submitted (13-MAR-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zfsh-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk

COMMENT

On Feb 24, 2003 this sequence version replaced gi:28460278.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zfsh-help@sanger.ac.uk

During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest, except on the rare
occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/projects/C.elegans/wormpep Repeat names
beginning 'Dr' were identified by The Recon repeat discovery system
(Zhirong Bao and Sean Eddy, submitted), and those beginning 'drr'
were identified by Rick Waterman (Stephen Johnson lab, WashU). For
further information see http://www/projects/d_rerio/fishmask.shtml
CH211-202P1 is from a CHORI-211 BAC library
VECTOR: pTARBAC2.1.

FEATURES

Location/Qualifiers
1..188317
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="CH211-202P1"
/clone_lib="CHORI-211"

ORIGIN

Query Match 12.1%; Score 49; DB 5; Length 188317;
Best Local Similarity 59.9%; Pred. No. 0.025;
Matches 82; Conservative 0; Mismatches 55; Indels 0; Gaps 0;
QY 1 ATGGGAATTTATCTCTCGGAAATCTAGTTGCCTGGGAGAGATCAAAAGAGCCCAAG 60
DB 120237 AAGACAATAACCACTCGATACACTGAGTTGTGAGAGAGCATCTCTAAAGCACAC 120178
QY 61 AAAGGAACCCAGATCAGAGAAAGAAACGGCAGGAAATGACTATATTGAAAGAAACTT 120
DB 120177 ACATGAACCTTGAGCGGATGGCTACAGCGCAAGACCACTCTGTGGCTAAGAA 120118
QY 121 CAAGATCAAGATAAGAA 137
DB 120117 CAGGAACAGGCAAGAA 120101

RESULT 15

AK125521/c AK125521 4218 bp mRNA linear PRI 09-SEP-2003
LOCUS Homo sapiens CDNA FLJ43533 fis, clone PLACE7001544.
DEFINITION
ACCESSION AK125521
VERSION AK125521.1 GI:34531645
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS
1 Tashiro, H., Yamazaki, M., Watanabe, K., Kumeagai, A., Itakura, S.,
Fukuzumi, Y., Fujimori, Y., Komiyama, M., Sugiyama, T., Irie, B.,
Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y.,
Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H.,
Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K.,
Wagatsuma, M., Murakawa, K., Kanehori, K., Takahashi-Fujii, A.,
Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S.,
Nagahara, K., Masuho, Y., Nagai, K. and Isogai, T.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 4218)
Isogai, T. and Yamamoto, J.
Direct Submission
TITLE
JOURNAL
COMMENT
Submitted (15-JUL-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7
Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI); (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.
FEATURES
Location/Qualifiers
source
1..4218
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PLACE7001544"
/tissue_type="Placenta"
/clone_lib="PLACE7"
/note="Cloning vector: pME18SFL3"

ORIGIN

Query Match 11.9%; Score 48; DB 9; Length 4218;
Best Local Similarity 52.0%; Pred. No. 0.075;
Matches 158; Conservative 0; Mismatches 140; Indels 6; Gaps 2;

QY 108 TGAAGAAGAACTTCAAGATCAAGATAGAAAGCCAGAGTTTCATCCACTTCTATCA 167
DB 3442 TGAATGTAAAGGTTCATGCGCAAGAGACCAATCTTGATGTTCTTTTGACAAGCA 3383
QY 168 GGAAACGAGAGATGCGAGTGTCTGAAGAGTGTCTACACTGTCAATTAATCA--CAT 224
DB 3382 GGCAATGTTGACAGAGCTTACTAGAGAGCTGTGCTTACCTCATCAATCATCGGGT 3323
QY 225 CCCCCATCAGAGATCCTCCCTGAGTCCCAATGATGAGGCTATGAGAACATGACTCCCT 284
DB 3322 TCTCTGTACAAGGCCATCAGGGAAGTCTGCTGAGAGTACTATGAGAATGTTCCCTGCAA 3263
QY 285 CACAGGAAGTGAGACAGTTTAGAAGAGGTGAGAGAGAGATATGCCCTTCTTAGGAC 344
DB 3262 AGCTGAGAGACCCAGAGAGTCTTGGGAGGAACTGAGACTGAGTATTCACCTTCACATAT 3203
QY 345 TTCTGTTAGTGGCCTTG---TTCTGCAACCCATGAGCATGATTATGAAGTTGTGTTCC 401
DB 3202 GCCTTCTACAGACCCAGGCGATGCCGATGCCAGAGATGAATGAATGAACTTCTCATGCC 3143
QY 402 ACAC 405
DB 3142 TCAC 3139

Search completed: April 22, 2004, 15:25:21
Job time : 1793 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 22, 2004, 17:06:22 ; Search time 2521 Seconds
(without alignments)
1599.125 Million cell updates/sec

Title: US-10-005-907-2

Perfect score: 715

Sequence: 1 MGNLYRLKSLGKNGKPK.....SVSRPCSTHEHDYEVVFP 135

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-O=/cgn2 1/USPTO.spool.D/US10005907/runat 21042004 113754 11441/app query.fasta_1.327
-DB=EST -QFMT=fastap -SUFFIX=rat -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NOR=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10005907 @CGN 1.1 4237 @runat 21042004 113754 11441 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hcc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hcc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	715	100.0	527	10	BF242113
2	173.5	24.3	1045	10	BF797507
3	169.5	23.7	924	12	BG389415
4	169.5	23.7	987	9	AL560376
5	169.5	23.7	1039	12	BM479887
6	169.5	23.7	1095	12	BG106563
7	169.5	23.7	1124	10	BF795688
8	166.5	23.3	1000	12	BM904106
9	166.5	23.3	1070	12	BM456595
10	166.5	23.3	1193	12	BM455198
11	159.5	22.3	1085	12	BM455407
12	156.5	21.9	508	29	AY404586
13	156.5	21.9	683	10	AW963366
14	156.5	21.9	876	13	BQ214406
15	150.5	21.0	1041	10	BG024891
16	149	20.8	1165	12	BM458339
17	146.5	20.5	1059	12	BM459235
18	145	20.3	495	9	AA311038
19	142	19.9	526	14	CB286394
20	133.5	18.7	522	12	BG149093
21	128	17.9	825	14	CB320849
22	127	17.8	520	12	BG145241
23	127	17.8	524	12	BG276802
24	127	17.8	526	12	BG276546
25	125	17.5	661	10	B8629482
26	125	17.5	3317	11	AK036553
27	122	17.1	560	13	BU696543
28	119	16.6	884	13	EX390865
29	118.5	16.6	815	13	BU92945
30	116.5	16.3	998	12	BG176880
31	112	15.7	479	13	BU429114
32	110.5	15.5	539	12	BG276071
33	109.5	15.3	538	10	AW519547
34	108	15.1	1015	12	BM477520
35	105.5	14.8	1527	29	AG032114
36	104.5	14.6	502	29	AY404588
37	104	14.5	591	10	BF797042
38	103.5	14.5	927	12	BG177483
39	102.5	14.3	982	13	EX391573
40	101.5	14.2	288	10	AW480906
41	101.5	14.2	875	29	AG181241
42	100.5	14.1	508	29	CE686854
43	100	14.0	594	28	BH085248
44	99.5	13.9	1242	12	BI834393
45	99.5	13.9	1361	12	BG722475

ALIGNMENTS

RESULT 1
BF242113
LOCUS 601880401F1 NIH_MGC_55 Homo sapiens cDNA clone IMAGE:410906 5',
DEFINITION mRNA sequence.
ACCESSION BF242113.1 GI:11156040
VERSION BF242113
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 527)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: CLONETECH Laboratories, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LICM993 row: a column: 01
 High quality sequence stop: 518.

FEATURES
 source
 1. .527
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4109064"
 /tissue_type="from acute myelogenous leukemia"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_lib="NIH_MGC_55"
 /note="Organ: bone marrow; Vector: pDNR-LIB (Clontech); Site 1: Sfii (ggcgctcgcc); Site 2: Sfii (ggcattatggcc); Double-stranded cDNA was prepared from cell line RNA. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CAGCGCATATGGCC-3' and 3' adaptor sequence: 5'-ATTAGAGCGGCGCGCATG-dT(30)BN-3', (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.65 kb (range 0.9-4.0 kb). 14/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

ORIGIN
 Alignment Scores:
 Pred. No.: 6,666-63 Length: 527
 Score: 715.00 Matches: 135
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 10 Gaps: 0

US-10-005-907-2 (1-135) x BF242113 (1-527)

Qy 1 MetGlyAenTyLeuLeuArgLysLeuSerCysLeuGlyGluAenGlnLysProLys 20
 Db 45 ATGGGAATTTATCTCTCGAATACTGATGCTGGAGAGATCAAGAGACCCCAAG 104
 Qy 21 LysGlyAenProAaspGluArgLysArgGlnGluMetThrPheGluArgLysLeu 40
 Db 105 AAAGGAAACCCAGATGAGGAAAGAAACGGCAGGAATGACTACATTTGAAAGAAACTT 164
 Qy 41 GlnAenGlnAenLysSerGlnGluValSerThrSerAenGlnGluAenGluAen 60
 Db 165 CAAGATCAAGATAGAGAAAGCCAGAGATTTTCATCCACTTTATTCAGAAACAGAAAT 224
 Qy 61 GlySerGlySerGluGluValCysTyThrValLeuAenHisPheProHisGlnArgSer 80
 Db 225 GGCAGTGGTTCTGAAGAAGTGTGTACACTGTCTATTAATCATCATCCATCCCATCAGATCC 284
 Qy 81 SerLeuSerSerAenAaspGlyTyrcGluAenLeuSerLeuThrArgLysValArg 100
 Db 285 TCCCTGAGCTCCATGATGAGTGTATGAGACATTTGCTCCCTCAGAGAAAGTGAAGA 344
 Qy 101 GlnPheArgGluArgSerGluThrGluTyThrValLeuAenHisPheProHisGlnArgPro 120
 Db 345 CAGTTTAGAGAAAGGTTCAGAGACAGAAATATGCCCTTCTTAGGACTTCTGTAGTAGGCT 404
 Qy 121 CysSerCysThrHisGluHisAspTyrcGluValValPheProHis 135

Db 405 TGTCTCTGACCCCATGAGCATGATTATGAGTGTGTGTTTCCACAC 449

RESULT 2
 BF797507
 LOCUS BF797507
 DEFINITION 1045 bp mRNA linear EST 12-JAN-2001
 60257174F1 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:4340323 5',
 mRNA sequence.
 ACCESSION BF797507
 VERSION BF797507.1 GI:12102561
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1045)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Louis Staudt, M.D., Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LICM9952 row: d column: 20
 High quality sequence stop: 687.

FEATURES
 source
 1. .1045
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4340323"
 /tissue_type="lymphoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_85"
 /note="Organ: lymph; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.867 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."

ORIGIN
 Alignment Scores:
 Pred. No.: 1,988-07 Length: 1045
 Score: 173.50 Matches: 48
 Percent Similarity: 51.61% Conservative: 32
 Best Local Similarity: 30.97% Mismatches: 54
 Query Match: 24.27% Indels: 21
 DB: 10 Gaps: 5

US-10-005-907-2 (1-135) x BF797507 (1-1045)

Qy 1 MetGlyAenTyLeuLeuArgLysLeuSerCysLeuGlyGluAenGlnLysPro--- 19
 Db 103 ATGGGAATTTCTCTCTGAGAGAAACAGCGGCGAGACACTCAGAGATGCTTGG 162
 Qy 20 -----LysLysGlyAenProAaspGluArgLysArg----- 30
 Db 163 AATGTGAGAAATCAAAAGCCCAACAGAGAAACATCCAGATGCTGGGATCACCATATCGCT 222
 Qy 31 -----GlnGluMetThrThrPheGluArgLysLeuGlnAasp 42
 Db 223 GAAGGTGTTTCTGCTTCCATGGAAAAAATACTCATCTTTTGAAGAAGG---CAAGAT 279
 Qy 43 GlnAenLysLysSerGlnGluValSerThrSerAenGlnGluAenGlnArgSer 62
 Db 280 TCCCAAAACAAAATGAAGAATGTCTACTTCCCATCCAGCAATGTTTCCAGCAGACC 339
 Qy 63 GlySerGluGluValCysTyThrValLeuAenHis---IleProHisGlnArgSer 81

Db	Accession	Source	Organism	Reference	Authors	Title	Journal	Comment
Db	340	TACTCAGAGGAGCTGTGCTATACCTCCTCATTCAATCCGAGGTTCTCTGTACAGGCGCATCA	399					
Db	82	LeuSERsArnaAaBpAaBpClyTyTtCtUuEniLeaBpSeLeuThrArgUySValArgIn	101					
Db	400	GCGAACTGCTGTGAAGAGTACTATGAAGATGTTCCTCGAAGCTGAGAGACCCAGAGG	459					
Db	102	PheArgUaArgSerGluThrGluTyAlaLeuLeuArgThrSerValSerArgProCys	121					
Db	460	TCCTTGAGAGCACTGAGCTGAGTATTCATCTTACATATATGCTCTTCTACAGACCCAGG	519					
Db	122	SerCys---ThrHisGluHisAspArgTyrGluValValPheProHis	135					
Db	520	CATGCCCATCCCAAGAGATAAATATGAACTTTCATGCGCTCC	564					
RESULT 3								
LOCUS	BG389415							
DEFINITION	602414106FL NIH_MGC_92 Homo sapiens cDNA clone IMAGE:4522512 5',							
ACCESSION	BG389415							
VERSION	BG389415.1							
KEYWORDS	GI:13282861							
SOURCE	EST.							
ORGANISM	Homo sapiens (human)							
REFERENCE	Homo sapiens							
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;							
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.							
JOURNAL	1 (bases 1 to 924)							
COMMENT	NIH-MGC http://mgs.nci.nih.gov/ National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgaabp@remail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/BLNT at: http://image.llnl.gov Plate: LLNL10423 row: 1 column: 01 High quality sequence stop: 726. Location/Qualifiers 1..924 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:4522512" /tissue_type="embryonal carcinoma, cell line" /lab_host="DHIOB (phage-resistant)" /clone_lib="NIH_MGC_92" /note="Organ: testis; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; cloned unidirectionally; oligo-dT primed. Average insert size 2.5 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC library."							
FEATURES								
Source								
Alignment Scores:								
Pred. No.:	4.4e-07	Length:	924					
Score:	169.50	Matches:	47					
Percent Similarity:	51.61%	Conservative:	33					
Best Local Similarity:	30.32%	Mismatches:	54					
Query Match:	23.71%	Indels:	21					
DB:	12	Gaps:	5					
US-10-005-907-2 (1-135) x BG389415 (1-924)								
QY	1	MetGlyAerTtyLeuLeuArgYsLeuSerCySleuGlyGluAsnGlnTysPro---	19					
Db	96	ATGGAGAAATCTCTGCTGAGAGAAAACAGCGGACGACAGAACACTCAAGAGATGCTTG	155					
QY	20	---LysAllyGlyAsnProAaBpGluGluArgYsArg---	30					

Db	Accession	Version	Source	Organism	Reference	Authors	Title	Journal	Comment
Db	156	AA07GAGATGCAAGGCCCAACAGAGAACATCCAGATGCTGGGATGATCACATATGCT	215						
Qy	31	-----GInGluMetThrThrpheGluValArgLysLeuGlnAsp	42						
Db	216	GAAGGCTGTTTCTGCTCCATGAGAAATAATCTATTTTGAAGAGAG--CAAGAT	272						
Qy	43	GlnAspLysLysSerGlnGlnValSerSerThrSerangGlnGlnGlnGlnGlnSer	62						
Db	273	TCCCAAAACGAAATAGAAAGATCATCTTACTCCATCCAGGACAAATGTGACCAAGC	332						
Qy	63	GlySerGlnGlnValCysTyrThrValIleasnHis--IleProHisGlnArgSerSer	81						
Db	333	TACCAAGAGAGCTGCTACTTACCTCATCATATGAGGGTTCTCTGTACAGGCCATCA	392						
Qy	82	LeuSerSerAsnAspAspGlyTyrGlnAsnIleAspSerLeuThrArgLysValArgGln	101						
Db	393	GGGAACCTCTGCTGAGAGTACTNAGAAAGTTCCGCAAGGTGAGAGCCAGAGAG	452						
Qy	102	PheArgGluArgSerGlnThrGluTyrAlaLeuLeuArgThrSerValSerArgProCys	121						
Db	453	TCTTGTGAGAGAGCTGAGACTGAGTATCTCTTCTTACATATGCTCTTACAGACCCAG	512						
Qy	122	SerCys---ThrHisGluHisAspTyrGlnValValAlaPheProHis	135						
Db	513	CATGCCCGATCCCAAGAGATGAAATATGAACTTCTCATGCTCTCAC	557						
RESULT 4									
LOCUS	AL560376	987 bp	mRNA	linear	EST 31-MAY-2003				
DEFINITION	AL560376 Homo sapiens B CELLS (RAMOS CELL LINE) Homo sapiens cDNA								
ACCESSION	AL560376								
VERSION	AL560376.2	GI:31284507							
KEYWORDS	EST.								
SOURCE	Homo sapiens (human)								
ORGANISM	Homo sapiens								
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;								
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.								
TITLE	1. (bases 1 to 987)								
JOURNAL	Li,W.B., Gruber,C., Jessee,J. and Polayes,D.								
COMMENT	Full-length cDNA libraries and normalization								
	Unpublished (2001)								
	On Feb 15, 2001 this sequence version replaced gi:12906778.								
	Contact: Genoscope								
	Genoscope - Centre National de Sequencage								
	BP 191 91006 EVRY cedex - France								
	Email: seq@genoscope.cns.fr, Web : www.genoscope.cns.fr								
	Library was constructed by Life Technologies, a division of								
	Invitrogen. This sequence belongs to sequence cluster 1964.r For								
	more information about this cluster, see								
	http://www.genoscope.cns.fr/								
	cg1-bln/cluster.cgi?seq=CS0D0007CB08QPKcluster=1964.r. Contact :								
	Feng Liang Email : fliang@lifetech.com URL :								
	http://fulllength.invitrogen.com/Invitrogen Corporation 1600								
	Faraday Avenue Genoscope sequence ID : CS0D0007CB08QPK1.								
FEATURES	location/Qualifiers								
SOURCE	1. 987								
	/organism="Homo sapiens"								
	/mol_type="mRNA"								
	/db_xref="taxon:9606"								
	/clone="CS0D0007YD15"								
	/tissue_type="B CELLS (RAMOS CELL LINE)"								
	/cell_line="RAMOS CELL LINE"								
	/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE)"								
	/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed								
	with a NotI-oligo(dT) primer. Five prime and enriched								
	double-strand cDNA was digested with Not I and								

Score: 169.50 Matches: 47
Percent Similarity: 51.61% Conservative: 33
Best Local Similarity: 30.32% Mismatches: 54
Query Match: 23.71% Indels: 21
DB: 9 Gaps: 5

US-10-005-907-2 (1-135) x AL560376 (1-987)

QY 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysPro--- 19
DB 179 ATGGGAAATTCCTGCTGAGAGAAAACAGCGGACAGCAACACTCAAGAGATGCTTGG 238
QY 20 -----LysLysGlyAsnProAspGluGlyArg----- 30
DB 239 AATGTAGAAATCAAGACCCCAACAGAGAACATCCAGATGCTGGATCACCATTCCT 298
QY 31 -----GlnGluMetThrThrPheGluArgLysLeuGlnAsp 42
DB 299 GAAGGCTTTCTGCTTCCTCCATGAGAAAAAATACATCTTTTGAAGAGAG---CAAGAT 355
QY 43 GlnAspLysLysSerGlnGluValSerSerThrSerAsnGlnGluAsnGlySer 62
DB 356 TCCCAAAACGAAATGAAAGATGATCTACTCCATCCAGAGAAATGTAACAGACC 415
QY 63 GlySerGluGluValCysTyrThrValIleAsnHis---IleProHisGlnArgSerSer 81
DB 416 TACTCAGAGAGAGCTGCTGCTTACCTCATCACTACGAGTTCCTGTAAGAGCCATCA 475
QY 82 LeuSerSerAsnAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArgGln 101
DB 476 GGGAAATCTGCTGAGAGTACTATGAGATGTCTCCAGAGAGAGAGAGAGAGAG 535
QY 102 PheArgGluArgSerGlnGluValSerSerThrSerValSerArgProCys 121
DB 536 TCCCTGGAGAGAGAGCTGCTGCTTACCTCATCACTACGAGTTCCTGTAAGAGCCATCA 595
QY 122 SerCys---ThrHisGlnHisAspTyrGluValValPheProHis 135
DB 596 CATGCCGATCCCAAGAGATGATGAACTTCTCATGCTCTCAC 640

RESULT 5
LOCUS BM479887 1039 bp mRNA linear EST 05-FEB-2002
DEFINITION AGENCOURT 6464830 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:557364
ACCESSION BM479887
VERSION BM479887.1 GI:18528929
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1039)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Straubeberg, Ph.D.
Email: cga@pds-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM12350 Row: 1 Column: 05
High quality sequence stop: 665.
Location/Qualifiers
1. 1039
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:557364"

FEATURES

1. 1039
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:557364"

/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_92"
/note="Organ: testis; Vector: pCMV-Sport6; site 1: NotI;
site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC library."

ORIGIN

Alignment Scores:

Pred. No.: 5,02e-07 Length: 1039
Score: 169.50 Matches: 47
Percent Similarity: 51.61% Conservative: 33
Best Local Similarity: 30.32% Mismatches: 54
Query Match: 23.71% Indels: 21
DB: 12 Gaps: 5

US-10-005-907-2 (1-135) x BM479887 (1-1039)

QY 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysPro--- 19
DB 81 ATGGGAAATTCCTGCTGAGAGAAAACAGCGGACAGCAACACTCAAGAGATGCTTGG 140
QY 20 -----LysLysGlyAsnProAspGluGlyArg----- 30
DB 141 AATGTAGAAATCAAGACCCCAACAGAGAACATCCAGATGCTGGATCACCATTCCT 200
QY 31 -----GlnGluMetThrThrPheGluArgLysLeuGlnAsp 42
DB 201 GAAGGCTTTCTGCTTCCTCCATGAGAAAAAATACATCTTTTGAAGAGAG---CAAGAT 257
QY 43 GlnAspLysLysSerGlnGluValSerSerThrSerAsnGlnGluAsnGlySer 62
DB 258 TCCCAAAACGAAATGAAAGATGATCTACTCCATCCAGAGAAATGTAACAGACC 317
QY 63 GlySerGluGluValCysTyrThrValIleAsnHis---IleProHisGlnArgSerSer 81
DB 318 TACTCAGAGAGAGCTGCTGCTTACCTCATCACTACGAGTTCCTGTAAGAGCCATCA 377
QY 82 LeuSerSerAsnAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArgGln 101
DB 378 GGAATCTGCTGAGAGAGTACTATGAGATGTCTCCAGAGAGAGAGAGAGAGAG 437
QY 102 PheArgGluArgSerGlnGluValSerSerThrSerValSerArgProCys 121
DB 438 TCCCTGGAGAGAGAGCTGCTGCTTACCTCATCACTACGAGTTCCTGTAAGAGCCATCA 497
QY 122 SerCys---ThrHisGlnHisAspTyrGluValValPheProHis 135
DB 498 CATGCCGATCCCAAGAGATGATGAACTTCTCATGCTCTCAC 542

RESULT 6
LOCUS BG106563 1095 bp mRNA linear EST 30-JAN-2001
DEFINITION 602290373F1 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:4385178 5'
ACCESSION BG106563
VERSION BG106563.1 GI:12600409
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1095)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Straubeberg, Ph.D.
Email: cga@pds-remail.nih.gov
Tissue Procurement: Louis Straub, M.D., Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L1AM10066 row: a column: 19
High quality sequence stop: 641.
Location/Qualifiers

FEATURES

source

1. 1095

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:4385178"

/issue_type="lymphoma, cell line"

/lab_host="DH10B (phage-resistant)"

/note="Organ: lymph; Vector: PCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.867 kb. Library enriched for
full-length clones and constructed by Life Technologies.

Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.:	5.32e-07	Length:	1095
Score:	169.50	Matches:	47
Percent Similarity:	51.61%	Conservative:	33
Best Local Similarity:	30.32%	Mismatches:	54
Query Match:	23.71%	Indels:	21
DB:	12	Gaps:	5

US-10-005-907-2 (1-135) x BGI06563 (1-1095)

```
Oy 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysPro--- 19
Db 43 ATGGGAATTTCTGCTGAGAGAAACAGCGGACGACGAACTCAAGAGATGCTTGG 102
Oy 20 -----LysLysGlyAsnProAspGlnLysArg----- 30
Db 103 AATGTGAATGCAAGACCCCAACAGAGAACTCCAGATCGGAGTACCATATCGCT 162
Oy 31 -----GlnGluMetThrThrPheGluArgLysLeuGlnAsp 42
Db 163 GAAGGTTGTTCTGCTTCATGAAAAAAATCACTATTGTAAGAGG---CAAGAT 219
Oy 43 GlnAspLysSerGlnGluValSerSerThrSerAsnGlnGluAsnGlySer 62
Db 220 TCCCAAAAGAAATGAAGATGCACTCTCCATCCAGACAAATGTTGACGAC 279
Oy 63 GlySerGluGluValCysTyrThrValIleAsnHis---IleProHisGlnArgSer 81
Db 280 TACTCAGAGAGAGCTGTGATACCTCATCAATCGGGTTCTCTGTACAAAGCCATCA 339
Oy 82 LeuSerSerAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArgGln 101
Db 340 GGAAGACTCTGCTCAAGAGACTATGAAATTTCCCTCCGAAAGCAGAGACCCAGAG 399
Oy 102 PheArgGluArgSerGlnThrGluTyrAlaLeuLeuArgThrSerValSerArgProCys 121
Db 400 TCCTTGAGAGAGACTGAGACTGATTCACCTTCAACATATGCCCTTCTACAGACCCGAG 459
Oy 122 SerCys---ThrHisGluHisAspTyrGluValIlePheProHis 135
Db 460 CATGCCCATCCCCAGAGATGAATATGAACTTCTCATGCCCTCAC 504
```

RESULT 7
BF795688 1124 bp mRNA linear EST 12-JAN-2001
LOCUS 60259560F1 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:4342719 5',
DEFINITION mRNA sequence.
ACCESSION BF795688
VERSION BF795688.1 GI:12100659
KEYWORDS EST.
SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

1 (bases 1 to 1124)

NIH-MGC <http://mgi.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

COMMENT

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-remail.nih.gov

Tissue Procurement: Louis Staudt, M.D., Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: L1AM9958 row: h column: 16

High quality sequence stop: 715.
Location/Qualifiers

FEATURES

1. 1124

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:4342719"

/issue_type="lymphoma, cell line"

/lab_host="DH10B (phage-resistant)"

/clone_id="NIH_MGC_85"

/note="Organ: lymph; Vector: PCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.867 kb. Library enriched for
full-length clones and constructed by Life Technologies.

Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.:	5.48e-07	Length:	1124
Score:	169.50	Matches:	47
Percent Similarity:	51.61%	Conservative:	33
Best Local Similarity:	30.32%	Mismatches:	54
Query Match:	23.71%	Indels:	21
DB:	10	Gaps:	5

US-10-005-907-2 (1-135) x BF795688 (1-1124)

```
Oy 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysPro--- 19
Db 118 ATGGGAATTTCTGCTGAGAGAAACAGCGGACGACGAACTCAAGAGATGCTTGG 177
Oy 20 -----LysLysGlyAsnProAspGlnLysArg----- 30
Db 178 AATGTGAATGCAAGACCCCAACAGAGAACTCCAGATCGGAGTACCATATCGCT 237
Oy 31 -----GlnGluMetThrThrPheGluArgLysLeuGlnAsp 42
Db 238 GAAGGTTGTTCTGCTTCATGAAAAAAATCACTATTGTAAGAGG---CAAGAT 294
Oy 43 GlnAspLysSerGlnGluValSerSerThrSerAsnGlnGluAsnGlySer 62
Db 295 TCCCAAAAGAAATGAAGATGCACTCTCCATCCAGACAAATGTTGACGAC 354
Oy 63 GlySerGluGluValCysTyrThrValIleAsnHis---IleProHisGlnArgSer 81
Db 355 TACTCAGAGAGAGCTGTGATACCTCATCAATCGGGTTCTCTGTACAAAGCCATCA 414
Oy 82 LeuSerSerAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArgGln 101
Db 415 GGAAGACTCTGCTCAAGAGACTATGAAATTTCCCTCCGAAAGCAGAGACCCAGAG 474
Oy 102 PheArgGluArgSerGlnThrGluTyrAlaLeuLeuArgThrSerValSerArgProCys 121
Db 475 TCCTTGAGAGAGACTGAGACTGATTCACCTTCAACATATGCCCTTCTACAGACCCGAG 534
Oy 122 SerCys---ThrHisGluHisAspTyrGluValIlePheProHis 135
```



```

Db      196  CCTTGGAAATGAGAAATGCAAGCCCAACAGAGAACATCCAGATCGCTGGATCACCAT 255
QY      31  -----GlnGluMetThrThrPheGluArgLysLeu 40
Db      256  ATGCTGAAGGGTGTCTGCTCCATGCGAAATACTCATTTTGAAGAAGG--- 312
QY      41  GlnAspGlnAspLysSerGlnGluValSerThrSerAsnGlnGluAsn 60
Db      313  CAAGATTCCCAAAACGAAATGAAGAATGCTACTTCCATCCATCCAGGACAATGTTGAC 372
QY      61  GlySerGlySerGluGluValCysThrValIleAsnHis---IleProHisGlnArg 79
Db      373  CAGACCTTACTCAGAGGAGCTGTGCTATACCTCATCAATCATCGGTTCTCTTACAGG 432
QY      80  SerSerLeuSerSerAsnAspGlyTyrGluAsnIleAspSerLeuThrArgLysVal 99
Db      433  CCATCAGGAACTCTGCTGAAGAGTACTATGAGATGTTCCCTGCAAGCTGAGAGACCC 492
QY      100  ArgGlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArg 119
Db      493  AGAGAGTCCTTGGAGGAACCTGAGACTGAGTATTCATCTTACATATGCCCTTCTACAG 552
QY      120  ProCysSerCys---ThrHisGluHisAspTyrGluValValPheProHis 135
Db      553  CCCAGGATGCCCATCCCAAGAGATGATGAACTTCTCATGCTCTCAC 603

RESULT 10
BM455198 1193 bp mRNA linear EST 05-FEB-2002
LOCUS AGENCOURT_6405645 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:5500141
DEFINITION 5', mRNA sequence.
ACCESSION BM455198.1 GI:18504238
VERSION 1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1193)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL
COMMENT Contact: Robert Strausberg, Ph.D.
Tissue Procurement: Lou Staudt
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12134 row: j column: 14
High quality sequence stop: 592.
Location/Qualifiers
1..1193
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5500141"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_85"
/notes="Organ: lymph; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.867 Kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

FEATURES
source
1..1193
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5500141"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_85"
/notes="Organ: lymph; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.867 Kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

ORIGIN
Alignment Scores: 1.18e-06 Length: 1193
Pred. No.: 166.50 Matches: 48
Score:

```

```

Percent Similarity: 52.23% Conservative: 34
Best Local Similarity: 30.57% Mismatches: 52
Query Match: 23.29% Indels: 23
DB: 12 Gaps: 6

US-10-005-907-2 (1-135) x BM455198 (1-1193)

QY      1  MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyGlu-----AsnGlnLysLys 18
Db      65  ATGGGAAATTTCTGCTGAGAGAAACATTTTCAGGCGGCGAGAGAACACATCAAGAGATG 124
QY      19  Pro-----LysLysGlyAsnProAspGluGluArgLysArg----- 30
Db      125  CTTTGAATGTGAGATGCAAGTCCAAAGCCCCAAACAGAGAACATCCAGATCGTGGGATCACC 184
QY      31  -----GlnGluMetThrThrPheGluArgLysLeu 40
Db      185  ATGCTGAAGGGTGTCTGCTCCATGCGAAATACTCATTTTGAAGAAGG--- 241
QY      41  GlnAspGlnAspLysSerGlnGluValSerThrSerAsnGlnGluAsnGluAsn 60
Db      242  CAAGATTCCCAAAACGAAATGAAGAATGCTACTTCCATCCATCCAGGACAATGTTGAC 301
QY      61  GlySerGlySerGluGluValCysThrValIleAsnHis---IleProHisGlnArg 79
Db      302  CAGACCTTACTCAGAGGAGCTGTGCTATACCTCATCAATCATCGGTTCTCTGTACAAGG 361
QY      80  SerSerLeuSerSerAsnAspGlyTyrGluAsnIleAspSerLeuThrArgLysVal 99
Db      362  CCATCAGGAACTCTGCTGAAGAGTACTATGAGATGTTCCCTGCAAGCTGAGAGACCC 421
QY      100  ArgGlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArg 119
Db      422  AGAGAGTCCTTGGAGGAACCTGAGACTGAGTATTCATCTTACATATGCCCTTCTACAG 481
QY      120  ProCysSerCys---ThrHisGluHisAspTyrGluValValPheProHis 135
Db      482  CCCAGGATGCCCATCCCAAGAGATGATGAACTTCTCATGCTCTCAC 532

RESULT 11
BM455407
LOCUS AGENCOURT_6407334 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:5500391
DEFINITION 5', mRNA sequence.
ACCESSION BM455407
VERSION 1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1085)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-r@mail.nih.gov
Tissue Procurement: Lou Staudt
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12135 row: d column: 24
High quality sequence start: 19
High quality sequence stop: 760.
Location/Qualifiers
1..1085
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5500391"

FEATURES
source
1..1085
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5500391"

```



```
source
1. .683
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="MAGE resequences, MAGE"
/notes="Vector: pBluescriptSKm"

ORIGIN
Alignment Scores:
Pred. No.: 6.59e-06 Length: 683
Score: 156.50 Matches: 37
Percent Similarity: 60.75% Conservatives: 28
Best Local Similarity: 34.58% Mismatches: 39
Query Match: 21.89% Indels: 3
DB: 10 Gaps: 3

US-10-005-907-2 (1-135) x AW963366 (1-683)

QY 31 GlnGluMetThrThrPheGluArgLysLeuGlnAspGlnAspLysLysSerGlnGluVal 50
DB 57 AAAAAAATACCTCATTTTGAAGAGAGG---CAAGATTCCCAAAACGAAATGAAGAATG 113
QY 51 SerSerThrSerAsnGlnGluAsnGlySerGlySerGlnGluValCysTyrThr 70
DB 114 TCATCTACTCCCATCCAGGACATGTTGACACACCTACTCAGAGGAGCTGTCTATACC 173
QY 71 ValIleAsnHis---IleProHisGlnArgSerSerLeuSerSerAsnAspAspGlyTyr 89
DB 174 CTCATCAATCATCGGGTCTCTGTACAGGCCATCAGGAACTCTGCTGAAGAGTACTAT 233
QY 90 GluAsnIleAspSerLeuThrArgLysValArgGlnPheArgGluArgSerGluThrGlu 109
DB 234 GAGAATGTTCCCTCAAGCTGAAAGACCCAGAGAGTCTTGGAGGAAGTGAAGATGAG 293
QY 110 TyrAlaLeuLeuArgThrSerValSerArgProCysSerCys---ThrHisGluHisAsp 128
DB 294 TATTCATCTTCTACATATGCTTCTACAGACCCAGGATGCCGATCCCGAAGATGAA 353
QY 129 TyrGluValValPheProHis 135
DB 354 TATGAACCTTCTCATGCTCTCAC 374

RESULT 14
BQ214406 876 bp mRNA linear EST 02-MAY-2002
LOCUS AGENCOURT 7592611 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:6070069
5', mRNA sequence.
ACCESSION BQ214406
VERSION BQ214406.1 GI:20395806
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 876)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1353 row: m column: 14
High quality sequence stop: 678.
Location/Qualifiers
1. .876
/organism="Homo sapiens"
/mol_type="mRNA"

FEATURES
source
/db_xref="taxon:9606"
/clone_lib="NIH_MGC_92"
/notes="Organ: testis; Vector: pCMV-SPORT6; Site 1: NOL1;
Site 2: Sall; Cloned unidirectionally; oligo-dr primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

ORIGIN
Alignment Scores:
Pred. No.: 8.72e-06 Length: 876
Score: 156.50 Matches: 37
Percent Similarity: 60.75% Conservatives: 28
Best Local Similarity: 34.58% Mismatches: 39
Query Match: 21.89% Indels: 3
DB: 13 Gaps: 3

US-10-005-907-2 (1-135) x BQ214406 (1-876)

QY 31 GlnGluMetThrThrPheGluArgLysLeuGlnAspGlnAspLysLysSerGlnGluVal 50
DB 394 AAAAAAATACCTCATTTTGAAGAGAGG---CAAGATTCCCAAAACGAAATGAAGAATG 440
QY 51 SerSerThrSerAsnGlnGluAsnGlySerGlySerGlnGluValCysTyrThr 70
DB 441 TCATCTACTCCCATCCAGGACATGTTGACACACCTACTCAGAGGAGCTGTCTATACC 500
QY 71 ValIleAsnHis---IleProHisGlnArgSerSerLeuSerSerAsnAspAspGlyTyr 89
DB 501 CTCATCAATCATCGGGTCTCTGTACAGGCCATCAGGAACTCTGCTGAAGAGTACTAT 560
QY 90 GluAsnIleAspSerLeuThrArgLysValArgGlnPheArgGluArgSerGluThrGlu 109
DB 561 GAGAATGTTCCCTCAAGCTGAGAGACCCAGAGAGTCTTGGAGGAGTCTGAGACTGAG 620
QY 110 TyrAlaLeuLeuArgThrSerValSerArgProCysSerCys---ThrHisGluHisAsp 128
DB 621 TATTCATCTTCTACATATGCTTCTACAGACCCAGGATGCCGATCCCGAAGATGAA 680
QY 129 TyrGluValValPheProHis 135
DB 681 TATGAACCTTCTCATGCTCTCAC 701

RESULT 15
BQ24891 1041 bp mRNA linear EST 24-JAN-2001
LOCUS 602275814F1 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:436899 5',
mRNA sequence.
ACCESSION BQ24891
VERSION BQ24891.1 GI:12410931
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1041)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Louis Staudt, M.D., Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
CDNA Library Arrayed by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10010 row: k column: 04
High quality sequence start: 2
```


GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 22, 2004, 16:07:56 ; Search time 349 Seconds

(without alignments)
1643.286 Million cell updates/sec

Title: US-10-005-907-2

Perfect score: 715

Sequence: 1 MGNYLRLKLSCLGKQKPK.....SVSRPCSTHEHDYEVVFPF 135

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2/1/USPTO.spool_p/US10005907/runat_21042004_113753_11422/app_query.fasta_1.327
-DB=N Geneseq 29Jan04 -QMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS-bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS-human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFWT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10005907 -CGN_1_1_708 -runat_21042004_113753_11422 -NCPU=6 -ICPU=3
-NO MMAP -LARGESQUERY -NEG SCORES=0 -WAIT -DSELOCK=100 -LONGLOG
-DEVO TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOXT=7

Database : N Geneseq 29Jan04:*

- 1: Geneseq1980s:*
- 2: Geneseq1990s:*
- 3: Geneseq2000s:*
- 4: Geneseq2001as:*
- 5: Geneseq2001bs:*
- 6: Geneseq2002as:*
- 7: Geneseq2003as:*
- 8: Geneseq2003bs:*
- 9: Geneseq2003cs:*
- 10: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	715	100.0	667	9	ADC30760 Human nov
2	715	100.0	3762	6	ABN81319 Human mas
3	169.5	23.7	732	7	ACC59951 Human IRA
4	159	22.2	1643	6	ABL99791 Human sec
5	149	20.8	775	7	ABX05108 Human nov
6	98	13.7	1892	2	AAZ24890 Human sec
7	98	13.7	1892	7	ADA56130 Gene enco
8	98	13.7	1892	7	ADA39940 Human sec

9	98	13.7	1892	9	ADC73578	Human sec
10	98	13.7	1892	9	ADD37623	Human sec
11	96.5	13.5	584	9	ADE62022	Rat gene
12	96	13.4	2301	2	AAV20445	Human c-t
13	96	13.4	2301	3	AAA09303	Human Trk
14	93	13.0	1824	4	ABL27495	Drosophil
15	92	12.9	654	5	AAS71153	DNA encod
16	92	12.9	654	5	AAS75467	DNA encod
17	92	12.9	654	5	AAS75467	DNA encod
18	91.5	12.8	3182	5	ABX71302	Human kid
19	91.5	12.8	7488	6	ABL33112	Human imm
20	90	12.6	1671	4	AAH99858	Human pro
21	90	12.6	1899	2	AAH99858	Human pro
22	90	12.6	1912	3	AAS50927	Human pro
23	90	12.6	7992	6	ABK32895	DNA encod
24	90	12.6	8083	5	AAK70692	DNA encod
25	90	12.6	8257	8	ACA90154	Human CDN
26	90	12.6	8493	7	ACA03930	CDNA down
27	90	12.6	8503	4	AAI58253	Human pol
28	90	12.6	8503	8	ADB48220	Novel hum
29	90	12.6	8527	4	AAI60039	Human pol
30	89.5	12.5	1027	2	AAK84690	CBALD05
31	89.5	12.5	1027	6	ABK47580	CDNA enco
32	89.5	12.5	1071	4	ABA83301	Human sec
33	89.5	12.5	1071	8	ACH04802	Novel hum
34	89.5	12.5	1071	8	ACD44612	Human CDN
35	89.5	12.5	5695	3	AAC76938	Human ORF
36	89.5	12.5	9415	6	ABL32295	Human imm
37	89	12.4	864	7	ABX09875	N. mening
38	89	12.4	902	6	ABQ46432	Oligonuci
39	89	12.4	902	6	ABQ46432	Oligonuci
40	89	12.4	13574	3	AAH1529	N. mening
41	89	12.4	110000	3	AAF21613	Continuation (8 of
42	89	12.4	17325	3	AAF21613	Neisseria
43	88.5	12.4	3430	8	ACF05253	Eimeria m
44	88	12.3	7409	4	ABL13405	Abi13405 Drosophil
45	88	12.3	10215	4	ABL13404	Abi13404 Drosophil

ALIGNMENTS

RESULT 1

ADC30760

ID ADC30760 standard; cDNA; 667 BP.

XX

AC ADC30760;

XX

DT 18-DEC-2003 (first entry)

XX

DE Human novel cDNA sequence, SEQ ID NO:842.

XX

Human; diagnostic; drug screening; forensics; gene mapping;
biodiversity assessment; Parkinson's disease; Alzheimer's disease;
neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
ulcers; osteoporosis; autoimmune disease; cancer;
molecular weight marker; food supplement; antiparkinsonian; nootropic;
neuroprotective; antianemic; anticoagulant; thrombolytic; vulnary;
antitumor; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
gene therapy; chromosome 1; Gene; ss.

XX Homo sapiens.

XX WC2003029271-A2.

PN

XX 10-APR-2003.

PD

XX 24-SEP-2002; 2002WO-US030474.

PF

XX 24-SEP-2001; 2001US-0124631P.

PR

XX (HYSE-) HYSEQ INC.

PA

XX Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;

PI

Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
 Haley-Vicente D, Drmanac RT;

WPI; 2003-371981/35.

P-PSDB; ADC31731.

New polynucleotide and polypeptide useful for diagnosing, preventing or
 treating conditions such as neurodegenerative diseases, anemias, platelet
 disorders, wounds, ulcers, osteoporosis, autoimmune diseases or
 cancer.

Claim 1; SEQ ID NO 842; 1185pp; English.

The invention relates to 971 novel human cDNA sequences (ADC29919-
 ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The
 invention also relates to nucleic acid sequences over 98% identical with
 the novel human cDNAs. The invention additionally encompasses expression
 vectors and host cells comprising a nucleic acid of the invention; the
 recombinant production of a polypeptide of the invention; an antibody
 against a polypeptide of the invention; a method of detecting
 polynucleotides or polypeptides of the invention; and methods of
 identifying a compound which binds to a polypeptide of the invention. The
 invention further discloses methods of preventing, treating or
 ameliorating a medical condition; kits comprising polynucleotide probes
 and/or monoclonal antibodies for carrying out the methods of the
 invention; methods for the identification of compounds that modulate the
 expression or activity of the polynucleotide and/or polypeptide; and 767
 contig sequences corresponding to the cDNA sequences of the invention
 (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
 -ADC33394). The nucleic acids and polypeptides of the invention are
 useful in diagnostics, drug screening, forensics, gene mapping, in the
 identification of mutations responsible for genetic disorders or other
 traits, for assessing biodiversity, and in producing many other types of
 data and products dependent on DNA and amino acid sequences. They are
 also used for treating diseases such as Parkinson's disease, Alzheimer's
 disease and other neurodegenerative diseases, anaemia, platelet
 disorders, wounds, ulcers, osteoporosis, autoimmune diseases or
 cancer. The nucleic acids may also be used as hybridisation probes or
 primers, and in the recombinant production of a protein. The polypeptides
 are also useful in generating antibodies, as molecular weight markers,
 and as food supplements. The present sequence represents a specifically
 claimed human cDNA sequence of the invention. Note: The sequence data for
 this patent did not form part of the printed specification, but was
 obtained in electronic format directly from WIPO at
 ftp.wipo.int/pub/published_pct_sequences.

Sequence 667 BP; 216 A; 147 C; 154 G; 150 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 8,248-76 Length: 667
 Score: 715.00 Matches: 135
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0

US-10-005-907-2 (1-135) x ADC30760 (1-667)

QY 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLeuProlys 20
 DB 186 ATGGGAAATATCTCTCGGAAACTCACTTGGCTGGGAGAGAAATCAAAAGAACCCCAAG 245
 QY 21 LysGlyAsnProAspGlnGluArgLysArgGlnGluMetThrThrPheGluArgLysLeu 40
 DB 246 AAGGNAACCCAGATGAGGAAAGAAACCGGAGGAATGACTACATTTGAAGAATACTT 305
 QY 41 GlnAspGlnAspLysSerGlnGluValSerThrSerAsnGlnGluAsnGluAsn 60
 DB 306 CAAGATCAAGATAAGAAACCAAGAAAGTTTCATCCACTTCTTAATCAGGAAACCGAAT 365
 QY 61 GlySerGlySerGluGluValCysTyrThrValIleAsnHisIleProHisGlnArgSer 80
 DB 366 GGCAGTGGTCTGAGAGAGTGGTCACTGCTCATCTGCTCATCTCATCTCATCTCATCTCATCT 425

QY 81 SerLeuSerSerAsnAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArg 100
 DB 426 TCCCTGAGCTCCAAATGATGATGGCTATGAGAACATTCGACTCCCTCACAGGAAGTGA 485
 QY 101 GlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArgPro 120
 DB 486 CAGTTTAGAGAAAGGTGAGACAGAGATATGCCCTCTCTAGGACTTCTGTGTAGGCT 545
 QY 121 CysSerCysThrHisGluHisAspTyrGluValValPheProHis 135
 DB 546 TGTCTCTGCACCCATGAGCATGATTATGAGTGTGTGTTCACAC 590

RESULT 2

ABN81319

ID ABN81319 standard; cDNA; 3762 BP.

XX AC ABN81319;

XX 30-AUG-2002 (first entry)

XX Human mast cell related gene MCL SEQ ID NO 1.

XX Human; mast cell; MC; antiallergic; antiinflammatory; antiasthmatic;
 KW vasotrophic; dermatological; allergy; hypersensitivity; rhinitis; asthma;
 KW gene; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 25..432

FT /tag= a

FT /product= "MCL"

PN WO200246389-A2.

XX 13-JUN-2002.

XX 07-DEC-2001; 2001WO-US046180.

XX 08-DEC-2000; 2000US-0251835P.

PR 14-MAR-2001; 2001US-0275479P.

PR 28-MAR-2001; 2001US-0279115P.

PR 02-APR-2001; 2001US-0280143P.

XX (UNIO) UCB SA.

XX Nocka X, Pirozzi G, Einstein R;

XX WPI; 2002-508560/54.

XX P-PSDB; ABB77569.

XX Novel isolated nucleic acids that are differentially expressed in mast
 cells in patients with allergic hypersensitivity, encoding proteins
 associated with mast cell regranulation and allergic hypersensitivity.

Claim 1; Page 95-97; 119pp; English.

XX The invention relates to isolated nucleic acid (ABN81319-ABN81324),
 corresponding to genes differentially expressed in mast cells following
 activation or in patients with allergic hypersensitivity disease, (I)
 that encodes proteins (ABB77569-ABB77575) (II) or a protein fragment of
 (II) if at least 6 amino acids. (II) is useful for identifying binding
 partners. (I) or (II) is useful for diagnosing or treating a disease
 state (e.g. allergic hypersensitivity, seasonal rhinitis, asthma,
 urticaria or atopic dermatitis or mastocytosis) in a subject which
 involves determining the level of expression of (I) or (II). A computer
 system, comprising a database containing information identifying the
 expression level in a tissue or at least one mast cell of (I), is useful
 for presenting information to identify the relative expression level of
 (I). (II) is used as a marker to detect, diagnose or identify an allergic
 response in a patient. The protein can also serve as a target that
 modulate gene expression or activity and as an antigen to raise

CC polyclonal or monoclonal antibodies. (II) is useful for identifying
 CC agents that modulate expression of the protein or agents, such as
 CC agonists or antagonists. The agonists or antagonists are useful for
 CC modulating biological activity and function of (II) and thus are useful
 CC for alleviating disease conditions such as allergic hypersensitivity,
 CC seasonal rhinitis, asthma, urticaria, atopic dermatitis or mastocytosis
 XX
 XX Sequence 3762 BP; 1220 A; 672 C; 680 G; 1190 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 8,19e-75 Length: 3762
 Score: 715.00 Matches: 135
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-10-005-907-2 (1-135) x ABN81319 (1-3762)

Qy 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyLysGlnGlnLysProLys 20
 Db 25 ATGGGAAATATCTCTCGGAAACTCAGTTGCTGGGAGAGATCAAAAGAGCCCAAG 84
 Qy 21 LysGlyAsnProAspGluGluArgLysArgGlnGlnMetThrThrPheGluArgLysLeu 40
 Db 85 AAAGGAACCCAGATGAGGAAGAAACGACGAGGAATGACTACATTTGAAGAAACTT 144
 Qy 41 GlnAspGlnAspLysSerGlnGlnLysSerThrSerGlnGlnGlnGlnGlnGln 60
 Db 145 CAAGATCAAGATAGAAAGCAAGAGTTTCTATCCACTTCTAATCAGGAAACGAGAAAT 204
 Qy 61 GlySerGlySerGluGluValCysTyrThrValIleAsnHisIleProHisGlnArgSer 80
 Db 205 GGCAGTGGTCTTGAAGAGTGCTGACTGCTCATTTAATCATCCCTCCATCAGATCC 264
 Qy 81 SerLeuSerSerAsnAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArg 100
 Db 265 TCCCTGAGTCCATCATGATGCTGATGATGATGATGATGATGATGATGATGATGATGAT 324
 Qy 101 GlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArgPro 120
 Db 325 CAGTTTAGAAGAAAGGTGAGACAGATATGCTCTTCTAGGACTTCTGTAGTAGGCTT 384
 Qy 121 CysSerCysThrHisGluHisAspTyrGluValValPheProHis 135
 Db 385 TGTCTCTGACCCATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 429

RESULT 3

ACC59951
 ID ACC59951 standard; cDNA; 732 BP.

AC ACC59951;

DT 07-JUL-2003 (first entry)

DE Human IRAP-28 encoding cDNA SEQ ID NO:63.

KW Human; immune response associated protein; IRAP; anti-HIV; anti-allergic;
 KW anti-inflammatory; antianemic; antiparkinsonian; nootropic; antithyroid;
 KW anticonvulsant; antiarteriosclerotic; antidiabetic; immunosuppressive;
 KW cyostatic; hepatotropic; dermatological; antidiabetic; nephrotropic;
 KW angiot; thyromimetic; neuroprotective; osteopathic; antithyroid;
 KW antiparasitic; antihelminthic; antipsoriatic; uropathic; protozoacide;
 KW antirheumatic; haemostatic; antibacterial; virucide; ophthalmological;
 KW fungicide; gene therapy; immune system disorder; neurological disorder;
 KW developmental disorder; muscle disorder; cell proliferative disorder;
 KW infection; gene; ss.

OS Homo sapiens.

FX Key Location/Qualifiers
 FH 134..670
 FT CDS /*tag= a

FT /product= "IRAP-28"

XX WO2003025542-A2.

XX 27-MAR-2003.

PF 19-SEP-2002; 2002WO-US029979.

XX 21-SEP-2001; 2001US-0324034P.

PR 05-OCT-2001; 2001US-0327395P.

PR 12-OCT-2001; 2001US-0328923P.

PR 19-OCT-2001; 2001US-0342810P.

PR 09-NOV-2001; 2001US-0344468P.

PR 21-NOV-2001; 2001US-0332140P.

PR 07-DEC-2001; 2001US-0340282P.

PR 09-JAN-2002; 2002US-0347693P.

PR 20-FEB-2002; 2002US-0358279P.

PR 01-MAR-2002; 2002US-0361088P.

PR 15-MAR-2002; 2002US-0364494P.

PR 10-MAY-2002; 2002US-0379876P.

PR 11-JUN-2002; 2002US-0388180P.

XX (INCY-) INCYTE GENOMICS INC.

XX Ho A, Baughn MR, Becha SD, Burford N, Elliott VS, Emerling BM;

XX Forsythe LJ, Gorvad AE, Griffin JA, Hafalia AJA, Honchell CD;

XX Ison CH, Burrill JD, Blake JJ, Lal PG, Lee EA, Marquis JP;

XX Lehr-Mason PM, Lee S, Sprague WM, Swarnakar A, Tang YT, Tran B;

XX Tran UK, Bhatia U, Chawla NK, Warren BA, Zheng W, Xu Y, Yue H;

XX WPI; 2003-363161/34.

XX P-PSDB; ABR43232.

XX New human immune response associated proteins and polynucleotides, useful
 XX for diagnosing, treating or preventing immune system disorders, e.g. AIDS
 XX or anemia, cell proliferative disorders, e.g. cancer, or neurological
 XX disorders.

XX Claim 5; Page 208; 213pp; English.

XX ACC59924 to ACC59958 encode the human immune response associated proteins
 XX given in ABR43205 to ABR43239, designated IRAP-1 to IRAP-29 (I). (I) have
 XX anti-HIV, anti-allergic, anti-inflammatory, antianemic, antiparkinsonian,
 XX nootropic, anticonvulsant, antiarteriosclerotic, antidiabetic, antipruritic,
 XX immunosuppressive, antithyroid, cytostatic, hepatotropic, dermatological,
 XX antidiabetic, nephrotropic, thyromimetic, neuroprotective, osteopathic,
 XX antithyroid, antiparasitic, antihelminthic, antipsoriatic, uropathic,
 XX ophthalmological, antirheumatic, haemostatic, antibacterial, virucide,
 XX protozoacide and fungicide activities, and can be used in gene therapy.
 XX Human IRAP polynucleotides, agonists and antagonists are useful for
 XX diagnosing, treating or preventing disorders associated with aberrant
 XX expression of IRAP, such as immune system disorders (e.g. AIDS, asthma,
 XX allergies, autoimmune thyroiditis, contact dermatitis, Crohn's disease,
 XX diabetes mellitus, glomerulonephritis, Good pasture's syndrome, gout,
 XX Graves' disease, Hashimoto's thyroiditis, irritable bowel syndrome,
 XX multiple sclerosis, osteoarthritis, osteoporosis, pancreatitis, Reiter's
 XX syndrome, rheumatoid arthritis, Sjogren's syndrome, uveitis),
 XX neurological disorders (e.g. Alzheimer's disease, Parkinson's disease or
 XX epilepsy), developmental disorders (e.g. renal tubular acidosis, anaemia
 XX or mental retardation), muscle disorder (e.g. cardiomyopathy, myocardiitis
 XX or dermatomyositis), cell proliferative disorders (e.g. arteriosclerosis,
 XX atherosclerosis, cirrhosis, hepatitis, psoriasis, paroxysmal nocturnal
 XX haemoglobinuria, polycythaemia vera, psoriasis, primary
 XX thrombocytopaenia or cancer), or viral, bacterial, fungal, parasitic,
 XX protozoan or helminthic infections

XX Sequence 732 BP; 213 A; 178 C; 178 G; 163 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1,28e-10 Length: 732
 Score: 169.50 Matches: 47
 Percent Similarity: 51.61% Conservative: 33
 Best Local Similarity: 30.32% Mismatches: 54

```

Query Match: 23.71% Indels: 21
DB:          Gaps: 5
US-10-005-907-2 (1-135) x ACC59951 (1-732)
Qy 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLysPro--- 19
Db 134 ATGGGAATTCCTCTGCTGAGAGAAACAGCGCGGAGAGAAACACTCAAGAGATGCTTGG 193
Qy 20 -----LysLysGlyAsnProAspGluGluArgLysArg----- 30
Db 194 AATGTGAGATGCAAGCCCAACACAGAGAACATCCAGATGCTGGGATCACCATATCGCT 253
Qy 31 -----GlnGluMetThrThrPheGluArgLysLeuGlnAsp 42
Db 254 GAAGGGTGTTCCTCCCTCCATGGAAAAAATACTCAATTTTGAAGAAGG---CAAGAT 310
Qy 43 GlnAspLysLysSerGlnGluValSerSerThrSerAsnGlnGluAsnGlySer 62
Db 311 TCCCAAAACGAATGAAGATGTCATCTACTCCATCCAGGACATGTTGACCAACC 370
Qy 63 GlySerGluGluValCysTyrThrValIleAsnHis---IleProHisGlnArgSerSer 81
Db 371 TACTCAGAGAGCTGTGTATACCTCATCAATCATCGGTTCTCTGACAAGGCCATCA 430
Qy 82 LeuSerSerAsnAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArgGln 101
Db 431 GGGAACTCTGCTGAAGATGACTATGAGATGTTCCCTGCAAGCTGAGACCCAGAGAG 490
Qy 102 PheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArgProCys 121
Db 491 TCCTTGGGAGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 550
Qy 122 SerCys---ThrHisGluHisAspTyrGluValValPheProHis 135
Db 551 CATGCCCGATCCCAAGAGATGAATATGAACTTCTCTGCTGCTCAC 595

RESULT 4
ID ABL99791 standard; cDNA; 1643 BP.
XX ABL99791;
AC ABL99791;
XX
XX
XX 03-OCT-2002 (first entry)
XX Human secretory polynucleotide (sptm) 46.
XX Human; ss; gene; secretory protein; secretory polynucleotides; SPTM;
KW SPTM-related disease; somatic gene therapy; germline gene therapy;
KW severe combined immunodeficiency; intracellular parasite protection;
KW fungal parasite; protozoan parasite; cell proliferative disorder; cancer;
KW immune disorder; AIDS; neurological disorder; Parkinson's disease;
KW motor neuron disorder; demyelinating disease; multiple sclerosis;
KW meningitis; abscess; prion diseases; cerebral palsy;
KW neuroskeletal disorder; peripheral nervous system disorder;
KW dermatomyositis; polymyositis; myopathy; myasthenia gravis;
KW mental disorder; Tourette's syndrome.
XX
XX Homo sapiens.
XX
XX WO200220756-A2.
XX
XX 14-MAR-2002.
XX
XX 30-AUG-2001; 2001WO-US027297.
XX
XX 05-SEP-2000; 2000US-0229747P.
XX 05-SEP-2000; 2000US-0229748P.
XX 05-SEP-2000; 2000US-0229749P.
XX 05-SEP-2000; 2000US-0229750P.
XX 05-SEP-2000; 2000US-0229751P.
XX 05-SEP-2000; 2000US-0230016P.
XX 05-SEP-2000; 2000US-0230583P.

```

```

PR 06-SEP-2000; 2000US-0230505P.
PR 06-SEP-2000; 2000US-0230514P.
PR 06-SEP-2000; 2000US-0230515P.
PR 06-SEP-2000; 2000US-0230517P.
PR 06-SEP-2000; 2000US-0230518P.
PR 06-SEP-2000; 2000US-0230519P.
PR 06-SEP-2000; 2000US-0230525P.
PR 06-SEP-2000; 2000US-0230526P.
PR 06-SEP-2000; 2000US-0230597P.
PR 06-SEP-2000; 2000US-0230599P.
PR 06-SEP-2000; 2000US-0230610P.
PR 06-SEP-2000; 2000US-0230864P.
PR 06-SEP-2000; 2000US-0230865P.
PR 06-SEP-2000; 2000US-0230988P.
PR 06-SEP-2000; 2000US-0230989P.
PR 06-SEP-2000; 2000US-0230990P.
PR 07-SEP-2000; 2000US-0230996P.
PR 07-SEP-2000; 2000US-0230997P.
PR 07-SEP-2000; 2000US-0230951P.
PR 07-SEP-2000; 2000US-0231163P.
PR 07-SEP-2000; 2000US-0231832P.
XX
XX (INCY-) INCYTE GENOMICS INC.
PA
XX
XX Stuart J., Lincoln SS, Altus CM, Dufour GB, Chalup MS, Hillman JL;
PI Jones AL, Yu JY, Wright RJ, Gietzen D, Liu TF, Yap PE, Dahl CR;
PI Momiyama MG, Bradley DL, Rohatgi SD, Harris B, Roseberry AM;
PI Gerstin EH, Peralta CH, David MH, Panzer SR, Flores V, Daffo A;
PI Marwaha R, Chen AJ, Chang SC, Au AP, Imman RR;
XX
XX WPI; 2002-315658/35.
DR F-PSDB; ABB97794.
XX

```

Polynucleotide sequences encoding human secretory proteins useful for gene therapy of e.g. genetic deficiency disorders, cancers, and diseases caused by intracellular parasites.

Claim 1; Page 275-276; 585pp; English.

The invention comprises the amino acid and coding sequences of human secretory (SPTM) proteins. The SPTM DNA and amino acid sequences are useful for treating a disease or condition associated with the expression of functional SPTM. The SPTM DNA sequences are useful for somatic or germline gene therapy to correct a genetic deficiency (e.g. severe combined immunodeficiency). The SPTM DNA sequences are also useful in providing protection against intracellular parasites (e.g. fungal parasites and protozoan parasites). The SPTM DNA and protein sequences are also useful for diagnosing cell proliferative disorders (e.g. cancer, immune disorders (e.g. AIDS), neurological disorders (e.g. Parkinson's disease), motor neuron disorders, demyelinating diseases (e.g. multiple sclerosis), meningitis, abscesses, prion diseases, cerebral palsy, neuroskeletal disorders, peripheral nervous system disorders, dermatomyositis and polymyositis, myopathy, myasthenia gravis, and mental disorders (e.g. Tourette's syndrome). cDNA sequences ABL99746 - ABL99929 represent human secretory polynucleotides of the invention

Sequence 1643 BP; 413 A; 382 C; 384 G; 460 T; 0 U; 4 Other;

```

Alignment Scores:
Pred. No.: 6.72e-09 Length: 1643
Score: 159.00 Matches: 47
Percent Similarity: 50.64% Conservative: 32
Best Local Similarity: 30.13% Mismatches: 55
Query Match: 22.24% Indels: 22
DB: 6 Gaps: 6

```

US-10-005-907-2 (1-135) x ABL99791 (1-1643)

```

Qy 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLysPro--- 19
Db 1096 ATGGGAATTCCTCTGCTGAGAGAAACAGCGCGGAGAGAAACACTCAAGAGATGCTTGG 1155
Qy 20 -----LysLysGlyAsnProAspGluGluArgLysArg----- 30

```

Db 1156 AATGTGAGATGCAAGCCCAACAGAGAAATCCAGATGCTGGGATCACCATATCGCT 1215
 Qy 31 -----GlnGluMetThrThrPheGluArgLysLeuGlnAsp 42
 Db 1216 GAAGGGTGTTCCTCCATCGAAATAATACTCATTTTGAAGAAGG---CAAGAT 1272
 Qy 43 GlnAspLysLysSerGlnGluValSerSerThrSer---AsnGlnGluAsnGluAsnGly 61
 Db 1273 TCCAAACGAAATGAAGATGTCATCTACTCCATCCAGCAGGACAATGTGACCAG 1332
 Qy 62 SerGlySerGlnGluValCysThrValIleAsnHis---IleProHisGlnArgSer 80
 Db 1333 ACTTACTCAGAGAGCTGTGTATACCTCATCAATCATCGGTTCTCTGTACAGGCCA 1392
 Qy 81 SerLeuSerSerAnAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArg 100
 Db 1393 TCAGGGAACCTGCTGAAGAGTACTATGAGATGTTCCCTGCAAGCTGAGAGACCCAGA 1452
 Qy 101 GlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArgPro 120
 Db 1453 GAGTCCTTGGGAGGAAGCTGAGACTGAGTATTCATCTCTANATATGCTCTCTACAGACCCC 1512
 Qy 121 CysSerCys---ThrHisGluHisAspTyrGluValValPheProHis 135
 Db 1513 AGGCATGCCGATCCCAAGAGATGAATATGAACTTTTCATGCTCTCAC 1560
 RESULT 5
 ID ABX05108
 XX ABX05108 standard; cDNA; 775 BP.
 AC ABX05108;
 DT 17-JAN-2003 (first entry)
 DE Human novel polynucleotide #123.
 KW Human; gene; ss; genetic disorder; gene mapping; medical imaging; cancer;
 KW neurodegenerative disorder; lymphoid cell disorder; osteoporosis;
 KW Parkinson's disease; Alzheimer's disease; bone degenerative disorder;
 KW osteoarthritis; periodontal disease; liver fibrosis; viral infection;
 KW fungal infection; bacterial infection; autoimmune disease; diabetes;
 KW atopic dermatitis.
 XX Homo sapiens.
 XX WO200274961-A1.
 XX 26-SEP-2002.
 XX 14-MAR-2002; 2002WO-US005109.
 XX 15-MAR-2001; 2001US-00810173.
 XX (HYSE-) HYSEQ INC.
 XX Tang Yt, Zhou P, Goodrich R, Asundi V, Zhang J, Zhao QH, Ren F;
 XX Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;
 XX Wehrman T, Wang J, Wang D, Drmanac RT;
 XX WPI; 2003-040556/03.
 XX P-PSDB; ABU00030.
 XX New isolated polypeptides and polynucleotides, useful for preventing,
 XX treating or ameliorating medical conditions, such as cancer,
 XX neurodegenerative disorders, lymphoid cell disorders, bone degenerative
 XX disorders, and infections.
 XX Claim 1; SEQ ID NO 123; 235pp; English.
 XX The invention relates to human polynucleotides and the polypeptides they
 XX encode. The polynucleotides and polypeptides are useful in diagnostics,
 XX forensics, gene mapping, medical imaging, identification of mutations

CC responsible for genetic disorders or other traits, assessing biodiversity
 CC and producing many other types of data and products dependent on DNA and
 CC amino acid sequences. They are also useful for preventing, treating or
 CC ameliorating medical conditions, such as cancer, neurodegenerative
 CC disorders (e.g. Parkinson's disease, Alzheimer's disease), lymphoid cell
 CC disorders, osteoporosis, osteoarthritis, bone degenerative disorders,
 CC periodontal disease, liver fibrosis, infections (e.g. viral, fungal or
 CC bacterial) or autoimmune diseases (e.g. diabetes, atopic dermatitis).
 CC Sequences ABX04986-ABX05511 represent human polynucleotides of the
 CC invention. Note: The sequence data for this patent is not represented in
 CC the printed specification but is based on sequence information supplied
 CC by the European Patent Office

SQ Sequence 775 BP; 226 A; 202 C; 160 G; 187 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 3,88e-08 Length: 775
 Score: 149.00 Matches: 37
 Percent Similarity: 60.19% Conservative: 28
 Best Local Similarity: 34.26% Mismatches: 39
 Query Match: 20.84% Indels: 4
 DB: Gaps: 4

US-10-005-907-2 (1-135) x ABX05108 (1-775)

Qy 31 GlnGluMetThrThrPheGluArgLysLeuGlnAspGlnAspLysLysSerGlnGluVal 50
 Db 329 AAAAAAATACTCATTTTGAAGAAGG---CAAGATTCCTCAACGAAATGAAGAATG 385
 Qy 51 SerSerThrSer---AsnGlnGluAsnGluAsnGlySerGlySerGluGluValCysTyr 69
 Db 386 TCATCTACTCCATCCAGCAGGACAATGTTGACACAGACTACTCAGAGGAGCTGTGTAT 445
 Qy 70 ThrValIleAsnHis---IleProHisGlnArgSerSerLeuSerSerAsnAspAspGly 88
 Db 446 ACCCTCATCAATCATCGGTTCTCTGTACAGCCCATCAGGAACTCTGCTCAAGAGTAC 505
 Qy 89 TyrGluAsnIleAspSerLeuThrArgLysValArgGlnPheArgGluArgSerGluThr 108
 Db 506 TATGAGATGTTCCCTGCAAACTCAGAGACCCAGAGAGTCTCTGGGAGGAACTGAGACT 565
 Qy 109 GlnTyrAlaLeuLeuArgThrSerValSerArgProCysSerCys---ThrHisGluHis 127
 Db 566 GAGTATTCACTTCTACATATGCTTCTACAGACCCCGCATGCGCCGATCCCCAGAGAT 625
 Qy 128 AspTyrGluValValPheProHis 135
 Db 626 GATATGAACCTTCTCATGCTCTCAC 649

RESULT 6
 AAZ24890
 ID AAZ24890 standard; DNA; 1892 BP.

XX AAZ24890;
 AC AAZ24890;
 XX 02-DEC-1999 (first entry)
 DT Human secreted protein gene 80 clone HLMAY54.

XX Human; secreted protein; fusion protein; gene therapy; protein therapy;
 KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
 KW developmental abnormality; foetal deficiency; blood; allergy; renal; ds;
 KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
 KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
 KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
 KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
 KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.

OS Homo sapiens.
 XX WO9947540-A1.
 XX 23-SEP-1999.

```

XX PF 18-MAR-1999; 99WO-US005804.
XX PR 19-MAR-1998; 98US-0078563P.
XX PR 19-MAR-1998; 98US-0078566P.
XX PR 19-MAR-1998; 98US-0078573P.
XX PR 19-MAR-1998; 98US-0078574P.
XX PR 19-MAR-1998; 98US-0078576P.
XX PR 19-MAR-1998; 98US-0078577P.
XX PR 19-MAR-1998; 98US-0078578P.
XX PR 19-MAR-1998; 98US-0078579P.
XX PR 19-MAR-1998; 98US-0078581P.
XX PR 01-APR-1998; 98US-0080312P.
XX PR 01-APR-1998; 98US-0080313P.
XX PR 01-APR-1998; 98US-0080314P.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Ruben SM, Ni J, Rosen CA, Yu G, Young PE, Feng P, Soppet DR;
XX PI Wei Y, Endress CA, Duan RD, Kyaw H, Ebner R, Lafleur DW, Olsen HS;
XX PI Shi Y, Moore PA;
XX DR WPI; 1999-562050/47.
XX DR P-PSDB; AAY41387.
XX PT New isolated human genes, useful for diagnosis and treatment of e.g.
XX PT cancers, neurological disorders, immune diseases, inflammation or blood
XX PT disorders.
XX PS Claim 1; Page 346; 484pp; English.
XX CC This sequence represents a nucleic acid molecule which encodes a secreted
XX CC human protein. The gene number, and the clone it is derived from, are
XX CC detailed in the descriptor line. The gene can be used to generate fusion
XX CC proteins by linking to the gene to a human immunoglobulin FC portion
XX CC (e.g. AA224802) for increasing the stability of the fused protein as
XX CC compared to the human protein only. The invention relates to 95 novel
XX CC genes and their fragments (nucleic acid sequences: AA224811-224907; amino
XX CC acid sequences AAY41308-Y41404) which are useful for preventing, treating
XX CC or ameliorating medical conditions e.g. by protein or gene therapy. Also,
XX CC pathological conditions can be diagnosed by determining the amount of the
XX CC new polypeptides in a sample or by determining the presence of mutations
XX CC in the new polynucleotides. Specific uses are described for each of the
XX CC 95 polynucleotides, based on which tissues they are most highly expressed
XX CC in (see AA224811 for described uses)
XX SQ Sequence 1892 BP; 461 A; 565 C; 472 G; 394 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 0.156 Length: 1892
Score: 98.00 Matches: 36
Percent Similarity: 38.26% Conservative: 21
Best Local Similarity: 24.16% Mismatches: 62
Query Match: 13.71% Indels: 30
DB: 2 Gaps: 4

US-10-005-907-2 (1-135) x AA224890 (1-1892)
QY 5 LeuLeuArgLysLeuSerCysLeuGluAenGlnLysLysProLysLysGlyAsnPro 24
DB 572 CTCTTACATCTCTCTTGTCTTGGAGCCAGGAGCCAGGAGCCAGACACAGCAGGAG 631
QY 25 AspGluGluArgLysArgGlnGluMetThrThrPheGluArgLysLeuGlnAspGlnAsp 44
DB 632 CAAGGAGTGAGGACAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 691
QY 45 LysLysSerGlnGluValSerThrSerAsnGlnGlnAsnGlnAsnGlnSerGlySer 64
DB 692 CAGGAGACGACAGACAGACAGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 751
QY 65 GluGlu-----ValCysTyrThrValIleAsnHisIleProHisGlnArgSer 80
DB 752 AAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 811

```

```

QY 81 ---SerLeuSerSerAsnAspAspGlyTyr----- 89
DB 812 GAATCTCTATCTTCTTAACCTTCTCTTTGCTCCCGGGTACGAGAGTAGAGTCTACT 871
QY 90 -----GluAsnIleAspSerLeuThrArgLysValArgGlnPheArgGlu--- 104
DB 872 CCTATGATAATGGAGACATCCAGGAGCTCATTCATCAGCCCGGAGGAATAGATGAATG 931
QY 105 -----ArgSerGluThrGluTyrAlaLeuLeuArg 114
DB 932 AATGAATATATGATGAGAACTCTCTACTGGAGAAACCAAAACCTCGCAGCCTCTGCGAG 991
QY 115 ThrSerValSerArgProCysSerCys 123
DB 992 CTGCCCCACACAGAGCCCTTGCTGGTGC 1018

RESULT 7
ADA56130
ID ADA56130 standard; DNA; 1892 BP.
XX AC ADA56130;
XX DT 20-NOV-2003 (first entry)
XX DE Gene encoding human secreted protein #309.
XX KW immunosuppressive; antiinflammatory; antiasthmatic; antiallergic;
KW cytostatic; cerebroprotective; neuroprotective; nootropic;
KW cardiovascular; antiarteriosclerotic; gene therapy;
KW human secreted protein; immune disorder; inflammation;
KW respiratory bowel disease; cancer; CNS disorder; neurodegenerative disorders;
KW inflammatory disorder; cancer; nephritis; Crohn's disease; asthma; allergy;
KW multiple sclerosis; ischaemic brain injury; Parkinson's disease;
KW Alzheimer's disease; atherosclerosis; myocarditis; chromosome mapping;
KW triple helix formation; antisense gene therapy; forensic biology; ds;
KW gene.
XX OS Homo sapiens.
XX PN WO2002102994-A2.
XX PD 27-DEC-2002.
XX PF 19-MAR-2002; 2002WO-US008278.
XX PR 21-MAR-2001; 2001US-0277340P.
XX PR 19-JUL-2001; 2001US-0306171P.
XX PR 13-NOV-2001; 2001US-0331287P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX PA Rosen CA, Ruben SM;
XX PI WPI; 2003-167512/16.
XX DR P-PSDB; ADA57026.
XX FT New human secreted polypeptides and polynucleotides, useful for
FT diagnosing, treating or preventing e.g. immune disorders, inflammatory
FT conditions, respiratory disorders, cancers, CNS disorders, or
FT neurodegenerative disorders.
XX PS Claim 21; SEQ ID NO 319; 1754pp; English.
XX CC The invention relates to 592 new human secreted polypeptides useful for
XX CC diagnosing, treating or preventing e.g. immune disorders, inflammatory
XX CC conditions, respiratory disorders, cancers, CNS disorders, or
XX CC neurodegenerative disorders, or polypeptides comprising an amino acid
XX CC sequence at least 95% identical to the new sequences. The polypeptides,
XX CC antibodies or antibody fragments that bind to the polypeptides, nucleic
XX CC acids encoding the polypeptides, agonists or antagonists that binds to
XX CC the polypeptide, are useful in preparing diagnostic or pharmaceutical
XX CC compositions for diagnosing, treating or preventing an e.g. immune

```

disorders, inflammatory conditions (e.g. inflammatory bowel disease, nephritis or Crohn's disease), respiratory disorders (e.g. asthma and allergy), cancers (e.g. gastric, ovarian or lung cancer), CNS disorders (e.g. multiple sclerosis or ischaemic brain injury), neurodegenerative disorders (e.g. Parkinson's disease or Alzheimer's disease), and cardiovascular disorders (e.g. atherosclerosis or myocarditis). The polynucleotides are useful for chromosome identification, chromosome mapping, for controlling gene expression through triple helix formation or antisense DNA or RNA, in gene therapy, for identifying individuals from minute biological samples, in forensic biology, and as hybridization probes. The polypeptides are useful for as molecular weight markers on sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) gels, to raise antibodies, for testing biological activities, and for treating or preventing neural disorders, immune system disorders, muscular, reproductive, gastrointestinal, pulmonary, cardiovascular, renal, proliferative and/or cancerous diseases. This sequence corresponds to a gene encoding one of the polypeptide of the invention. Note: The sequence data for this patent did form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 1892 BP; 461 A; 565 C; 472 G; 394 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 0.156 Length: 1892
 Score: 98.00 Matches: 36
 Percent Similarity: 38.26% Conservative: 21
 Best Local Similarity: 24.16% Mismatches: 52
 Query Match: 13.71% Indels: 30
 DB: 7 Gaps: 4

US-10-005-907-2 (1-135) x ADA56130 (1-1892)

QY 5 LeuLeuArgLysLeuSerCysLeuGluAsnGlnLysLysProLysLysGlyAsnPro 24
 Db 572 CTCTCAATCTCTTCTGTCCTCCGGAAGCCAGGACGCGCCAGACACAGCAGGAG 631
 QY 25 AspGluGluArgLysArgGlnGluMetThrThrPheGluArgLysLeuGlnAspGlnAsp 44
 Db 632 CAAGGAGTGAGCAGCAGGAGGAGGAGCCGACACAGACAGCAGGAGGCGCAGAA 691
 QY 45 LysLysSerGlnLysValSerSerThrSerAsnGlnGluAsnGlnLysGlySer 64
 Db 692 CAGGAGAGCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 751
 QY 65 GluGlu-----ValCysThrThrValIleAsnHisIleProHisGlnArgSer 80
 Db 752 AAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 811
 QY 81 ---SerLeuSerSerAsnAspGlyTyr----- 89
 Db 812 GAATCTCTATCTTCTTACCTCTCTTTGCTCCCGGTACGAGAGTAGAGTCTACT 871
 QY 90 -----GluAsnIleAspSerLeuThrArgLysValArgGlnPheArgGlu--- 104
 Db 872 CCTATGATATGAGAACATCCAGGAGCTCATTCGATCAGCCAGGAAATAGATGAATG 931
 QY 105 -----ArgSerGluThrGluThrAlaLeuLeuArg 114
 Db 932 AATGAATATATGATGAGAACTCTACTGGAGAAACCAAACCTGGCAGCTCTCTGAG 991
 QY 115 ThrSerValSerArgProCysSerCys 123
 Db 992 CTGCCCCACAGAGCCTTGCTGTGC 1018

RESULT 8

ADA39940

ID ADA39940 standard; cDNA; 1892 BP.

XX AC ADA39940;

XX DT 20-NOV-2003 (first entry)

XX

DE Human secreted protein encoding cDNA.

XX Human; secreted protein; cancer; hyperproliferative disorder;
 KW rheumatoid arthritis; autoimmune disorder; haematopoietic disorder;
 KW anaemia; allergic reaction; asthma; cardiovascular disorder;
 KW wound healing; cytostatic; immunosuppressive; neurotropic; neuroprotective;
 KW antiviral; antiallergic; hepatotropic; antidiabetic; antiinflammatory;
 KW vulnary; candiant; gene therapy; ss.

OS Homo sapiens.

XX WO2002102993-A2.

PN 27-DEC-2002.

XX 19-MAR-2002; 2002WO-US008123.

XX 21-MAR-2001; 2001US-0277340P.

XX 19-JUL-2001; 2001US-0306171P.

XX 13-NOV-2001; 2001US-0331287P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM;

XX WPI; 2003-175238/17.

XX New human secreted proteins and nucleic acid molecules, useful for
 preparing a diagnostic or pharmaceutical composition for diagnosing,
 preventing or treating cancer or other hyperproliferative disorder,
 asthma, allergies or AIDS.

PS Claim 9; SEQ ID NO 322; 3205pp; English.

XX The invention relates to novel genes ADA39629-ADA40565 and proteins
 ADA40566-ADA41501 for human secreted proteins, useful for preventing,
 treating or ameliorating medical conditions e.g. by protein or gene
 therapy. The polypeptides, nucleic acid molecules, antibodies or their
 fragments, and agonists or antagonists that bind to the polypeptide are
 useful for preparing a diagnostic or pharmaceutical composition for
 diagnosing or treating cancer or other hyperproliferative disorder. The
 polypeptides and nucleic acid molecules are also useful for detecting,
 preventing, diagnosing, prognosticating, treating or ameliorating cancer
 or other hyperproliferative disorders including neoplasms, autoimmune
 disorders (e.g. diabetes, rheumatoid arthritis, systemic lupus
 erythematosus, multiple sclerosis, autoimmune thyroiditis or haemolytic
 anaemia), haematopoietic or haematological disorders (e.g. anaemia,
 thrombocytopenia), allergic reactions including asthma or eczema,
 inflammatory disorders (e.g. ischaemia-reperfusion injury, inflammatory
 bowel disease or Crohn's disease), neurodegenerative disorders (e.g.
 Alzheimer's disease or Parkinson's disease), cardiovascular disorders
 (e.g. atherosclerosis, myocarditis), infectious diseases (bacterial,
 fungal or viral infections including HIV/AIDS), or wound healing and
 disorders of epithelial cell proliferation. The nucleic acids are also
 useful for chromosome identification, radiation hybrid mapping or long-
 range restriction mapping, as molecular weight markers, or as
 hybridization or diagnostic probes. The polypeptides and antibodies are
 useful for providing immunological probes for differential identification
 of the tissues immunohistochemistry assays. Note: The sequence data for
 this patent did not form part of the printed specification, but was
 obtained in electronic format directly from WIPO at
 ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1892 BP; 461 A; 565 C; 472 G; 394 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 0.156 Length: 1892
 Score: 98.00 Matches: 36
 Percent Similarity: 38.26% Conservative: 21
 Best Local Similarity: 24.16% Mismatches: 62
 Query Match: 13.71% Indels: 30
 DB: 7 Gaps: 4

Alignment Scores:

Pred. No.: 0.587 Length: 1824
 Score: 93.00 Matches: 36
 Percent Similarity: 37.20% Conservatives: 25
 Best Local Similarity: 21.95% Mismatches: 51
 Query Match: 13.01% Indels: 52
 DB: 4 Gaps: 5
 US-10-005-907-2 (1-135) x ABL27495 (1-1824)
 QY 3 AsnTyrLeuLeuArgLysLeuSerCys---LeuGlyGluAenGlnLysLysProLysLys 21
 Db 940 AACGAGGATTGCGGATCTGTGTCTCTGACACGACCGCGGAGAGGACGCAAG 999
 QY 22 -----GlyAsnProAspGluGluArgLysArgGlnGluMet 33
 Db 1000 CTGGCGGGGAATGCGAAGATTGCGGATACACGCCGCGGTGTCGCGAGGAGTG 1059
 QY 34 ThrThrPheGluArgLysLeuGlnAspGlnAspLysLysSerGlnGluVal----- 50
 Db 1060 GCCGCGTACCAAGAACTACGTGATGATGATGATGATGATGATGATGATGATGAT 1119
 QY 50 ----- 50
 Db 1120 AATCTAGAGTCAAGAGCTGCTGCTTACTTGACGAGGAGCGTCCCATGTGGCGGCA 1179
 QY 51 -----SerSerThrSerAenGlnGluAenGluAenGlySer 62
 Db 1180 AATCGCTGTGCGCGGTGCGGAGCAGCCACTAGGAATGCGCTTCGATGATGCGGCGAT 1239
 QY 63 GlySerGlu-----GluValCysTyrThrValIleAenHisIleProHis 77
 Db 1240 GTTCCAGTCCACACGCAATGCGATGACGACCATCCCGCTTGAGGAACATGCGAG 1299
 QY 78 GlnArgSer-----SerLeuSerSerAenAspAspGlyTyrGlu 90
 Db 1300 CAGCGGCAACTACCCAGGATCTACGTGATGCGCCACACGCTGAACGACGAGCGTTGAG 1359
 QY 91 AsnIleAspSerLeuThrArgLysValArgGlnPheArgGluArgSerGluThrGluTyr 110
 Db 1360 TATGTGGCTCAGTGGAGCGCGGATCCAGCACTGAGGAGGAGGCGCACCGCCACG 1419
 QY 111 AlaLeuLeuArg 114
 Db 1420 GCCCATCTACAA 1431

RESULT 15

AAS71153
 ID AAS71153 standard; cDNA; 654 BP.
 AC AAS71153;
 XX
 DT 13-FEB-2002 (first entry)
 DE DNA encoding novel human diagnostic protein #5957.
 XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
 XX Homo sapiens.
 OS
 XX
 PN W0200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US008631.
 XX
 PR 31-MAR-2000; 2000US-00540217.
 PR 23-AUG-2000; 2000US-00649167.
 XX
 XX (HYSE-) HYSEQ INC.
 PA Drmanac RT, Liu C, Tang YT;
 XX
 PI

XX

DR WPI; 2001-639362/73.

DR P-PSDB; ABG06966.

XX

PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX

PS Claim 1; SEQ ID NO 6957; 103pp; English.

XX

CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
 CC coding sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX

SQ Sequence 654 BP; 304 A; 81 C; 181 G; 88 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 0.198 Length: 654
 Score: 92.00 Matches: 35
 Percent Similarity: 41.01% Conservatives: 22
 Best Local Similarity: 25.18% Mismatches: 53
 Query Match: 12.87% Indels: 30
 DB: 5 Gaps: 3

US-10-005-907-2 (1-135) x AAS71153 (1-654)

QY 14 GluAenGlnLysLysProLysGlyGlnAsnProAspGluGluArgLysArgGlnGluMet 33
 Db 169 GAAAGGAA 228
 QY 34 ThrThrPheGluArgLysLeuGlnAspGlnAspLysLysSerGlnGluValSerSerThr 53
 Db 229 GAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGA 288
 QY 54 SerAenGlnGluAenGluAenGlySerGlySerGluGluValCysTyrThrValIleAen 73
 Db 289 GAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGA 330
 QY 74 HisIleProHisGlnArgSerSerLeuSer----- 83
 Db 331 GAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGA 390
 QY 84 SerAenAspAspGlyTyrGlnAenIleAspSerLeuThrArg----- 97
 Db 391 TCAGACATCAACGGATTGAAAAA-ATTACTTCCAGTACACCTATTTTAGCAGTACTAGA 449
 QY 98 -----LysValArgGlnPheArgGluArgSerGluThrGluTyr 110
 Db 450 GGATGCGTTTAGCAGTCTCTGGAAAAACACGACATAAAGGCAAGAAGAAGAAGAAGA 509
 QY 111 AlaLeuLeuArgThrSerValSerArgProCysSerCysThrHisGluHisAspTyr 129
 Db 510 GAAACAGGGGATTCAGATGATCGAGAGCAAAAGGATGTTCCACAGATGATGACTAT 566

Search completed: April 22, 2004, 17:20:13
Job time : 360 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 22, 2004, 17:57:48 ; Search time 334 Seconds
(without alignments)
1822.349 Million cell updates/sec

Title: US-10-005-907-2

Perfect score: 715

Sequence: 1 MGYLRLKSLCLGENQKPK.....SVSRPCSTHEHDYEVFFH 135

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Zgapop 6.0 , Zgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2907579 seqs, 2254313464 residues

Total number of hits satisfying chosen parameters: 5815158

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool_p/US10005907/runat_21042004_113756_11532/app_query.fasta_1.327
-DB=Published Applications NA -QFMT=fastap -SURF1=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10005907@cgn2_1.1.333 @runat_21042004_113756_11532
-NCPU=6 -ICPU=1 -NO MMAP -LARGOQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/2/pubpna/US07_PUB.seq:
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:
17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:
18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:
19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
------------	-------	--------------	----	-------------

1	715	100.0	3762	15	US-10-005-907-1	Sequence 1, Appli
2	98	13.7	1892	13	US-10-653-595-90	Sequence 90, Appl
3	98	13.7	1892	13	US-03-397-945-90	Sequence 30, Appl
4	96	13.4	2301	13	US-10-242-943-3	Sequence 3, Appli
5	93	13.0	1903	13	US-10-424-599-8569	Sequence 8569, Ap
6	92	12.9	6324	16	US-10-115-831-54	Sequence 54, Appl
7	91.5	12.8	2862	13	US-10-263-929-22	Sequence 22, Appl
8	91.5	12.8	5109	13	US-10-342-887-1552	Sequence 1552, Ap
9	91.5	12.8	5109	13	US-10-172-118-1552	Sequence 1552, Ap
10	91.5	12.8	5261	10	US-03-814-353-19577	Sequence 19577, A
11	91.5	12.8	7488	15	US-10-311-455-1085	Sequence 1085, Ap
12	90.5	12.7	651	13	US-10-261-175A-9	Sequence 9, Appli
13	90	12.6	1173	13	US-10-027-632-10372	Sequence 10372, A
14	90	12.6	1173	16	US-10-027-632-10372	Sequence 10372, A
15	90	12.6	1671	13	US-10-296-115-693	Sequence 693, App
16	90	12.6	1899	10	US-03-746-783-77	Sequence 77, Appl
17	90	12.6	7992	10	US-09-893-519A-140	Sequence 140, App
18	90	12.6	8491	15	US-10-133-013-260	Sequence 260, App
19	90	12.6	8493	14	US-10-071-766-51	Sequence 51, Appl
20	90	12.6	8503	15	US-10-037-270-130	Sequence 130, App
21	90	12.6	8503	16	US-10-117-722-130	Sequence 130, App
22	89.5	12.5	1071	13	US-09-933-767-118	Sequence 118, App
23	89.5	12.5	1071	13	US-10-004-860-118	Sequence 118, App
24	89.5	12.5	1071	15	US-10-023-282-118	Sequence 118, App
25	89.5	12.5	1107	13	US-10-424-599-112896	Sequence 112896,
26	89.5	12.5	9415	15	US-10-311-455-268	Sequence 268, App
27	89	12.4	864	16	US-10-320-800-49	Sequence 49, Appl
28	88	12.3	99934	16	US-10-085-117-73	Sequence 73, Appl
29	87	12.2	3720	14	US-10-007-805-576	Sequence 576, App
30	87	12.2	3720	15	US-10-076-622-576	Sequence 576, App
31	87	12.2	3720	15	US-10-124-805-576	Sequence 576, App
32	87	12.2	123192	15	US-10-175-523-71	Sequence 71, Appl
33	86.5	12.1	776	13	US-10-424-599-23224	Sequence 23224, A
34	86.5	12.1	1181	13	US-10-415-134-3	Sequence 3, Appli
35	86.5	12.1	1302	9	US-09-879-957-17	Sequence 17, Appl
36	86.5	12.1	7095	13	US-10-205-331-67	Sequence 67, Appl
37	86	12.0	1632	16	US-10-085-117-54	Sequence 54, Appl
38	86	12.0	1725	15	US-10-198-846-10857	Sequence 10857, A
39	86	12.0	1886	15	US-10-262-666-41	Sequence 41, Appl
40	86	12.0	1886	16	US-10-085-117-53	Sequence 53, Appl
41	86	12.0	3883	15	US-10-177-293-331	Sequence 331, App
42	85.5	12.0	1482	9	US-09-938-842A-585	Sequence 585, App
43	85.5	12.0	1482	11	US-09-938-842A-585	Sequence 585, App
44	85.5	12.0	2257	13	US-10-424-599-141192	Sequence 141192,
45	85.5	12.0	2264	13	US-10-424-599-13803	Sequence 13803, A

ALIGNMENTS

RESULT 1
US-10-005-907-1
; Sequence 1, Application US/10005907
; Publication No. US20030166881A1
; GENERAL INFORMATION:
; APPLICANT: Union Chimique Belge, S.A.
; APPLICANT: No. US20030166881A1A, Karl
; APPLICANT: Pitozzi, Gregory
; APPLICANT: Einstein, Richard
; TITLE OF INVENTION: NOVEL GENES ASSOCIATED WITH ALLERGIC HYPERSENSITIVITY AND MAINT CEI
; FILE REFERENCE: 053529-5005
; CURRENT APPLICATION NUMBER: US/10/005.907
; CURRENT FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 3762
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (25)..(432)

```

; OTHER INFORMATION:
US-10-005-907-1
Alignment Scores:
Pred. No.: 2,31e-85 Length: 3762
Score: 715.00 Matches: 135
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 15 Gaps: 0

US-10-005-907-2 (1-135) x US-10-005-907-1 (1-3762)
QY 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGluAsnGlnLysLysProLys 20
Db 25 ATGGGAATATTCTCTCGAAATCTCAGTTGCTGGGAGAGAAATCAAAAGAAAGCCCAAG 84
QY 21 LysGlyAsnProAspGluGluArgLysArgGlnGluMetThrThrPheGluArgLysLeu 40
Db 85 AAAGGAACCCAGATGAGGAAGAAAGAAACGGCAGGAATGACTACATTGGAAGAAACTT 144
QY 41 GlnAspGlnAspLysSerGlnGluValSerThrSerAsnGlnGluAsnGln 60
Db 145 CAAGATCAAGATAAGAAAGCCAGAAAGTTTCATCCACTTCTAATCAGGAAAGCGAGAT 204
QY 61 GlySerGlySerGluGluValCysTyrThrValIleAsnHisIleProHisGlnArgSer 80
Db 205 GGCATGGTTCTGAGAGAGTGTGTACTGCTATTATCATCTCCCTCCATCAGAGATCC 264
QY 81 SerLeuSerSerAsnAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArg 100
Db 265 TCCCTGAGCTCCAATGATGCTGCTATGAGAACATGACTCCCTCACAAGGAAAGTGAGA 324
QY 101 GlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArgPro 120
Db 325 CAGTTTAGAGAAAGTGTAGACAGACAGAAATGCCCCCTCTTAGACTTCGTAGAGCCCT 384
QY 121 CysSerCysThrHisGluHisAspTyrGluValValPheProHis 135
Db 385 TGTTCCTGCACCCATGAGCATGATTATGAAGTTGTGTTCACAC 429

RESULT 2
US-10-653-595-90
; Sequence 90, Application US/10653595
; Publication No. US20040048304A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et. al.
; TITLE OF INVENTION: 95 Human secreted proteins
; FILE REFERENCE: P2027PL1
; CURRENT APPLICATION NUMBER: US/10/653,595
; PRIOR FILING DATE: 2003-09-03
; PRIOR APPLICATION NUMBER: US 09/397945
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: PCT/US99/05804
; PRIOR FILING DATE: 1999-03-18
; PRIOR APPLICATION NUMBER: 60/078,566
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 60/078,576
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 60/078,573
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 60/078,574
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 60/078,579
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 60/080,314
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080,312
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/078,578
; PRIOR FILING DATE: 1998-03-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 470

; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 90
; LENGTH: 1892
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-653-595-90
Alignment Scores:
Pred. No.: 0.0109 Length: 1892
Score: 98.00 Matches: 36
Percent Similarity: 38.26% Conservative: 21
Best Local Similarity: 24.16% Mismatches: 62
Query Match: 13.71% Indels: 30
DB: 13 Gaps: 4

US-10-005-907-2 (1-135) x US-10-653-595-90 (1-1892)
QY 5 LeuLeuArgLysLeuSerCysLeuGluGluAsnGlnLysLysProLysLysGlyAsnPro 24
Db 572 CTCTACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 631
QY 25 AspGluGluArgLysArgGlnGluMetThrThrPheGluArgLysLeuGlnAspGlnAsp 44
Db 632 CAAGGAGTGGAGCACACAGGAGGAGCGCACACAGAAACACACAGCAGCAGAGGGGCGAGAA 691
QY 45 LysLysSerGlnGluValSerThrSerAsnGlnGluAsnGlnGluAsnGlySerGlySer 64
Db 692 CAGGAAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 751
QY 65 GluGlu-----ValCysTyrThrValIleAsnHisIleProHisGlnArgSer 80
Db 752 AAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 811
QY 81 ---SerLeuSerSerAsnAspAspGlyTyr----- 89
Db 812 GAATCTCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 871
QY 90 -----GluAsnIleAspSerLeuThrArgLysValArgGlnPheArgGlu--- 104
Db 872 CCTATGATAATGAGAACATCCAGGAGCTCATTCGATCAGCCCGAGAAATAGATGAATG 931
QY 105 -----ArgSerGluThrGluTyrAlaLeuLeuArg 114
Db 932 AATGAATATATGATGAGAACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 991
QY 115 ThrSerValSerArgProCysSerCys 123
Db 992 CTGCCCCACACAGAGCTTGTGTGTC 1018

RESULT 3
US-09-397-945-90
; Sequence 90, Application US/09397945
; Publication No. US20030065139A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc. et al.
; TITLE OF INVENTION: 95 Human secreted proteins
; FILE REFERENCE: P2027P1
; CURRENT APPLICATION NUMBER: US/09/397,945
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: PCT/US99/05804
; PRIOR FILING DATE: 1999-03-18
; PRIOR APPLICATION NUMBER: 60/078,566
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 60/078,576
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 60/078,573
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 60/078,574
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 60/078,579
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 60/080,314
; PRIOR FILING DATE: 1998-04-01
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 470

```


US-10-424-599-8569

Alignment Scores:

Pred. No.: 0.0518 Length: 1903
Score: 93.00 Matches: 36
Percent Similarity: 43.33% Conservative: 16
Best Local Similarity: 30.00% Mismatches: 52
Query Match: 13.01% Indels: 16
DB: 13 Gaps: 5

US-10-005-907-2 (1-135) x US-10-424-599-8569 (1-1903)

QY 3 AsnTyrLeuLeuArgLysLeuSerCysLeu---GlyGluAsnGln----- 16
DB 710 AACAGTCTGATTTCAGAGATGACACTACTAGAGGAGAAAACGAACTGGTAAATGAAG 769
QY 17 -----LysLysProLysLysGlyAsnProAspGluGluArgLysArgGlnGluMet 33
DB 770 TTTCTGAACAGAAACATCATAGGCTCAATCAGATTCAAGTATCAGTCAAGTGAAGGA 829
QY 34 ThrThrPheGluArgLysLeuGlnAspGlnAspLysLysSerGlnGluValSerThr 53
DB 830 GAAATGAGAGAGAGAGAGATCGCTATGAGGACAAATAACAGGAAAGATCACCCAGT 889
QY 54 SerAsnGlnGluAsnGluAsnGlySerGly---SerGluGluValCysTyrThrValIle 72
DB 890 CACCAACAGACACAAAGAAATGGCAAGACTACAAAGAGATGCTGTGACAGAAATAT 949
QY 73 AsnHisIleProHisGlnArgSerLeuSerSerAsn---AspAspGlyTyrGluAsn 91
DB 950 AACAGTCCAGTCCAGAAAGAGGTGTCCAAAAGGGTCAATTGGATTTCAGGGTATGAATCA 1009
QY 92 IleAspSerLeuThrArgLysValArgGlnPhe-----ArgGluArgSer 106
DB 1010 AGTGAAGAGAAAGAGAGAGAAAGAAATGACTATGAGGCAAAAATGCGGGAAGATCA 1069

RESULT 6

US-10-115-831-54
; Sequence 54, Application US/10115831
; Publication No. US20030219743A1

; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyao
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 792CIP2ADIV
; CURRENT APPLICATION NUMBER: US/10/115,831
; PRIOR FILING DATE: 2002-04-02
; PRIOR APPLICATION NUMBER: 09/667,298
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 09/577,408
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 178
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 54
; LENGTH: 6324
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3573)
US-10-115-831-54

Alignment Scores:

Pred. No.: 0.392 Length: 6324
Score: 92.00 Matches: 27
Percent Similarity: 38.60% Conservative: 17
Best Local Similarity: 23.68% Mismatches: 50
Query Match: 12.87% Indels: 20
DB: 16 Gaps: 3

US-10-005-907-2 (1-135) x US-10-115-831-54 (1-6324)

QY 14 GluAsnGlnLysLysProLysLysGlyAsnProAspGluGluArgLysArgGlnGluMet 33
DB 337 GACAATGACGAGCCAGGCGGATGAGGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 396
QY 34 ThrThrPheGluArgLysLeuGlnAspGlnAspLysLysSerGlnGluValSerThr 53
DB 397 GAGGAGATCGAGGAGGAGGATGAGGACGATCAGGAGGATGAGGAGGATGAGGAGGATGA 456
QY 54 SerAsnGlnGluAsnGluAsnGlySerGlySerGluGluVal----- 67
DB 457 GAAGAGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 516
QY 68 -----CysTyrThrValIleAsnHisIleProHisGlnArgSerSerLeuSerSerAsn 85
DB 517 ATGAATTGTCAAT-----ACTCGAATAATGCAAGACACAGAAAGGATGATAACAAT 570
QY 86 AspAspGlyTyrGluAsnIleAspSerLeuThrArgLysVal----- 99
DB 571 AATGACGAATATGACAAATACGATGAATGCTGTGGCCAAAGTCAATTGTTAAACCTCGGCAA 630
QY 100 -----ArgGlnPheArgGluArgSerGluThrGlu 109
DB 631 ATCGCTGAGGATGACGCTACCGGCCAGGACTGAGTCAGAA 672

RESULT 7

US-10-263-929-22
; Sequence 22, Application US/10263929
; Publication No. US20040067335A1

; GENERAL INFORMATION:
; APPLICANT: Kim, Jaeseob
; APPLICANT: Galant, Ron
; TITLE OF INVENTION: Alzheimer's Disease Linked Genes
; FILE REFERENCE: LSD-07417
; CURRENT APPLICATION NUMBER: US/10/263,929
; CURRENT FILING DATE: 2002-10-03
; NUMBER OF SEQ ID NOS: 213
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 22
; LENGTH: 2862
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-263-929-22

Alignment Scores:

Pred. No.: 0.148 Length: 2862
Score: 91.50 Matches: 32
Percent Similarity: 47.90% Conservative: 25
Best Local Similarity: 26.89% Mismatches: 45
Query Match: 12.80% Indels: 17
DB: 13 Gaps: 4

US-10-005-907-2 (1-135) x US-10-263-929-22 (1-2862)

QY 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLysProLys 20
DB 394 ATCGGCCAGTCGTTGTTGAAGAGAAACAGACCTAACCGAGAGGAAACGAGCTGCTGGAG 453
QY 21 LysGlyAsnProAspGluGluArgLysArgGlnGluMetThrThrPheGluArgLysLeu 40
DB 454 GAG-----CAGGTGGAACACATCAGGAGGAGGTTCTCAGCTCCGGCATGAGCTG 504
QY 41 GlnAspGlnAspLysSerGlnGluValSerThrSerAsnGlnGluAsnGluAsn 60
DB 505 TCCATGAAGATGAGCTGCTTCACTTACACCGAGGCTCGGAGGAG----- 552
QY 61 GlySerGlySerGluGluValCysTyrThrThrValIleAsnHisIleProHisGlnArgSer 80
DB 553 ---AGTGAGCCCGAGTCCGTTTGTCTCAACC-----CCGTTGAAGAGGAAT 594
QY 81 SerLeuSerSerAsnAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArg 100

```

; Sequence 1552, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1552
; LENGTH: 5109
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM_014965
; DATABASE ENTRY DATE: 2001-06-18
; US-10-172-118-1552

Alignment Scores:
Pred. No.: 0.338      Length: 5109
Score: 91.50         Matches: 32
Percent Similarity: 47.90%      Conservative: 25
Best Local Similarity: 26.89%    Mismatches: 45
Query Match: 12.80%             Indels: 17
DB: 13                     Gaps: 4

US-10-005-907-2 (1-135) x US-10-172-118-1552 (1-5109)

QY 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLysProLys 20
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 610 ATCGGCCAGTCGTTGTTGAAGAAGAACCAAGACCCTAACCGAGGAGACGAGCTGCTGGAG 669

QY 21 LysGlyAsnProAspGluGluArgLysArgGlnGluMetThrThrPheGluArgLysLeu 40
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 670 GAG-----CAGGTGGAACACATCAGGAGGAGGAGGTCTCAGCTCCGGCATGAGCTG 720

QY 41 GlnAspGlnAspLysLysSerGlnGluValSerThrSerAsnGlnGlnGluAsnGluAsn 60
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 721 TCCATGAAGATGAGCTGCTTCAGTTCTACACAGCGCTGCGGAGAG-----768

QY 61 GlySerGlySerGluGluValCysTyrThrValIleAsnHisIleProHisGlnArgSer 80
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 769 ---AGTGAGCCGAGTCCGTTTGTCTCAACC-----CCGTTGAAGAGGAAT 810

QY 81 SerLeuSerSerAsnAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArg 100
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 811 GAGTCGCTCTCTCAGTCCAGAAATTAATTTTCAATTTGGATTCTCTTCAAAGAGAGCTGAAA 870

QY 101 GlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArg 119
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 871 GACCTTGAAGAG-----GAGAAATGTTGACTTCGATCCGAGGCCAGCCAG 915

RESULT 10
US-09-814-353-19577
; Sequence 19577, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; US-10-172-118-1552
```

```

; Sequence 1552, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1552
; LENGTH: 5109
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-342-887-1552

Alignment Scores:
Pred. No.: 0.338      Length: 5109
Score: 91.50         Matches: 32
Percent Similarity: 47.90%      Conservative: 25
Best Local Similarity: 26.89%    Mismatches: 45
Query Match: 12.80%             Indels: 17
DB: 13                     Gaps: 4

US-10-005-907-2 (1-135) x US-10-342-887-1552 (1-5109)

QY 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLysProLys 20
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 610 ATCGGCCAGTCGTTGTTGAAGAAGAACCAAGACCCTAACCGAGGAGACGAGCTGCTGGAG 669

QY 21 LysGlyAsnProAspGluGluArgLysArgGlnGluMetThrThrPheGluArgLysLeu 40
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 670 GAG-----CAGGTGGAACACATCAGGAGGAGGAGGTCTCAGCTCCGGCATGAGCTG 720

QY 41 GlnAspGlnAspLysLysSerGlnGluValSerThrSerAsnGlnGlnGluAsnGluAsn 60
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 721 TCCATGAAGATGAGCTGCTTCAGTTCTACACAGCGCTGCGGAGAG-----768

QY 61 GlySerGlySerGluGluValCysTyrThrValIleAsnHisIleProHisGlnArgSer 80
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 769 ---AGTGAGCCGAGTCCGTTTGTCTCAACC-----CCGTTGAAGAGGAAT 810

QY 81 SerLeuSerSerAsnAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArg 100
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 811 GAGTCGCTCTCTCAGTCCAGAAATTAATTTTCAATTTGGATTCTCTTCAAAGAGAGCTGAAA 870

QY 101 GlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArg 119
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 871 GACCTTGAAGAG-----GAGAAATGTTGACTTCGATCCGAGGCCAGCCAG 915

RESULT 9
US-10-172-118-1552
```

```

; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1085
; LENGTH: 7488
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-1085

Alignment Scores:
Pred. No.: 0.583 Length: 7488
Score: 91.50 Matches: 29
Percent Similarity: 44.22% Conservative: 36
Best Local Similarity: 19.73% Mismatches: 51
Query Match: 12.80% Indels: 31
DB: 15 Gaps: 5

US-10-005-907-2 (1-135) x US-10-311-455-1085 (1-7488)

QY 14 GluAsnGlnLysLys-----Pro 19
DB 6080 GAATAATCTAATAAATAACGACTTCTCCGGCTCGACAAATACACTACTACATAA 6021

QY 20 LysLysGlyAsnProAspGluGluArgLysArgGlnGlu-----MetThrThr 35
DB 6020 AAACAAATAACCCCGATATAAAAAAAAAAAAAATAAACAAACCCCTAACTCTCTCATACTAAC 5961

QY 36 PheGluArg-----LysLeuGlnAspGlnAspLysLysSer 47
DB 5960 TTATAAATCTAACCCAAATCTTAACACCTTTAAATTAACGAAAAAACTTAAAAATCT 5901

QY 48 GlnGluValSerThrSerAsnGlnGlnAsnGluAsnGlySerGlySerGluGluVal 67
DB 5900 AAAACCCCTTACCAACGAAATACCCCTTAAAAATAACCAAAAAATAAACCTACTCGCCCTT 5841

QY 68 CysTyrThr---ValIleAsnHisIleProHisGlnArgSerSerLeuSerSerAsnAsp 86
DB 5840 CPAATTCCTTCCTACTCAACTATCTCGATACCAACGTAATAAATAAAAAATAAATAAAC 5781

QY 87 AspGlyTyrGluAsnIleAspSerLeuThrArgLysValArgGlnPheArgGluArgSer 106
DB 5780 CACAATATCTATAAAATTCATAATATTTTAAAAAATCGCTTACTAAAAACAACGAA 5721

QY 107 GluThrGluTyrAlaLeuLeuArgThrSerValSerArgProCysSerCysThrHisGlu 126
DB 5720 AAAACGCCACCAACCACTCGAAAA-----AAACCAATAATAACACACGTTTC 5673

QY 127 HisAspTyrGluValValPhe 133
DB 5672 CACGATAAAAAAAATCTAT 5652

RESULT 12
US-10-261-175A-9
; Sequence 9, Application US/10261175A
; Publication No. US20040038222A1
; GENERAL INFORMATION:
; APPLICANT: DIETRICH, WILLIAM
; APPLICANT: WATTERS, JAMES W.
; TITLE OF INVENTION: ANTHRAX SUSCEPTIBILITY GENE
; FILE REFERENCE: 56491(71250)
; CURRENT APPLICATION NUMBER: US/10/261,175A
; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: 60/325,864
; PRIOR FILING DATE: 2001-09-29
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: Patent In Ver. 2.1

```

; SEQ ID NO 9
; LENGTH: 651
; TYPE: DNA
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Kiflc nucleic
; OTHER INFORMATION: acid sequence
US-10-261-175A-9

Alignment Scores:
Pred. No.: 0.0243 Length: 651
Score: 90.50 Matches: 39
Percent Similarity: 36.97% Conservative: 22
Best Local Similarity: 23.64% Mismatches: 51
Query Match: 12.66% Indels: 53
DB: 13 Gaps: 7

US-10-005-907-2 (1-135) x US-10-261-175A-9 (1-651)

Qy 16 GlnLysLysPProLysLysGlyAsnProAspGluGluArgLysArgGlnGluMetThrThr 35
Db 4 GAGAAACCAACCCAGCCTCGTGTGACGAGGAGTGGCGCAAGGCCCGAGAGCGAGCGTCC 63
Qy 36 PheGluArgLysLeuGlnAspGlnAspLysLysSerGlnGluValSerSerThrSerAsn 55
Db 64 CGTCTGAGAAACTACAGAGCAAGACAAACAG---CAGAAAGTGGAGTTTCGGAAAGG 120
Qy 56 GlnGluAsnGlu---AsnGlySerGlySerGluGluValCysTyrThr 70
Db 121 ATGGAGAAAGAGGTGCTGATTTCATCCAGGACAGTGGACAGGTCAAGAAAGTTTCAG 180
Qy 71 ValIleAsnHisIleProHisGlnArgSer----- 80
Db 181 CCTATGACAAAGATA-----GAGCGGAGCATACTACATGATGTGTAGAGGTGGCTGGC 234
Qy 81 -----SerLeuSerSerAsnAsp----- 87
Db 235 CTCACATCTTCTCTCTGGAGAGATGATGATCTGCTGCTATGTCATGATCTTCAAAAG 294
Qy 88 -----GlyTyrGluAsnIleAspSerLeu----- 95
Db 295 GAGTTTGACCCCTCAGATGAAGAGCTAGACTCTCCTACCGCTCATGAGAGGAGTGGGACCC 354
Qy 96 -----ThrArgLysValArgGlnPheArgGluArgSerGluThrGluTyr 110
Db 355 CAGAAGGCTGAGGAGAACCGGAAGCTAAGAGGCTGCTCAGAGCAGGAGGAAGGCA 414
Qy 111 AlaLeuLeuArgThrSerValSerArgProCysSerCysThrHisGluHisAspTyrGlu 130
Db 415 GCGCAGCAGGACCTCGCGTTGTGAGTCTCTGCCAGC-----GACTACAAG 459
Qy 131 ValValPheProHis 135
Db 460 GACAAGTATAGCCAT 474

RESULT 13

US-10-027-632-10372
; Sequence 10372, Application US/10027632
; Publication No. US2002019871A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827,129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218

; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10372
; LENGTH: 1173
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-10372

Alignment Scores:
Pred. No.: 0.0658 Length: 1173
Score: 90.00 Matches: 38
Percent Similarity: 40.27% Conservative: 22
Best Local Similarity: 25.50% Mismatches: 55
Query Match: 12.59% Indels: 34
DB: 13 Gaps: 6

US-10-005-907-2 (1-135) x US-10-027-632-10372 (1-1173)

Qy 1 MetGlyAsnTyrLeuLeuArg-----LysLeuSerCysLeuGlyGlu 14
Db 542 ATGGGCAACTCCATCATGAGACCAATTTCTTCAGCTTCCAGTTTTCGCCCATGCTTAC 601
Qy 15 AsnGlnLysLysProLysLysGlyAsnProAsp-----GluGluArgLys 29
Db 602 CCTCAGATGAAGCCGAGTCACTCACAATTCGTGTCTAGTTAACTTACAGAGATCC 661
Qy 30 ArgGlnGluMetThrThrPheGluArgLysLeuGlnAspGlnAspLysLysSerGlnGlu 49
Db 662 CGTTAVGGAATGACTGACATGACCAATTTCTTCGTGACCTCGAAACTAACAGCATGCTA 721
Qy 50 ValSerSerThrSerAsnGlnGlu-----AsnGluAsnGlySerGlySer 64
Db 722 AATGACTCTAGTCTGATGAAGAGTTAAATGAATAGACAGTGAAGATGGT----- 772
Qy 65 GluGluValCysTyrThrValIleAsnHisIleProHisGlnArgSerSerLeuSerSer 84
Db 773 -----TTAAGTGTATGGATCACCAGACATCAGGCTGCTGCA 811
Qy 85 -----AsnAspAspGlyTyrGluAsnIleAspSerSerLeuThrArgLysValArg 100
Db 812 GAGCAGCTGATGGCTCAGATGGCAACAATTTATTGGAGACCAAGGGGATTCATTAGA 871
Qy 101 GlnPheArgGluArgSerGluThrGluTyr-----AlaLeuLeuArgThrSer 116
Db 872 AGATTCAATGATAGTTCCAGTGGCCCTTTCTCTCTCCTCACCATGATCGACGTAGC 931
Qy 117 ValSerArgProCysSerCysThrHis 125
Db 932 ATCTCTCGTCACATAGAAACATCCAC 958

RESULT 14

US-10-027-632-10372
; Sequence 10372, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827,129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483

```
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10372
; LENGTH: 1173
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-10372

Alignment Scores:
Pred. No.: 0.0658 Length: 1173
Score: 90.00 Matches: 38
Percent Similarity: 40.27% Conservativeness: 22
Best Local Similarity: 25.50% Mismatches: 55
Query Match: 12.59% Indels: 34
DB: 16 Gaps: 6

US-10-005-907-2 (1-135) x US-10-027-632-10372 (1-1173)
QY 1 MetGlyAenTyrLeuLeuArg-----LysLeuSerCysLeuGlyGlu 14
DB 542 ATGGCACTCCATCATGACGAGACCCCAATCTTCAGCTTCCAGTTTCGCCCATGCTTAC 601
QY 15 AenGlnLysLysProLysLysGlyAsnProAsp-----GluGluArgLys 29
DB 602 CCTCAGATGAAGCGCAAGTCACTCACAAATTCGTCTAGTTAACTTGACAGAGATCC 661
QY 30 ArgGlnGluMetThrThrPheGluArgLysLeuGlnAspGlnAspLysLysSerGlnGlu 49
DB 662 CGTAYGAAGTACTGACATGACCAATCTTCTGCTGACCTGGAACTAACAGATCCTA 721
QY 50 ValSerSerThrSerAsnGlnGlu-----AenGluAenGlySerGlySer 64
DB 722 AATGACTCTAGTTCTGTATGAAGAGTTAAATGAATAGACAGTGAAGATGGT----- 772
QY 65 GluGluValCysTyrThrValIleAsnHisIleProHisGlnArgSerLeuSer 84
DB 773 -----TTAGTCTAGGATCCACAGATCAGGCTGTCTGCA 811
QY -----AsnAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArg 100
DB 812 GAGCAGCTGATGGGCTCAGATGGCAACAAATATTGGAGACCAAGGGGATTCATTAGA 871
QY 101 GlnPheArgGluArgSerGluThrGluTyr-----AlaLeuLeuArgThrSer 116
DB 872 AGATTCTAGATAGTCTTCAGTGGCCCTTTTGCTCTTCTCACCATGATCGACGTAGC 931
QY 117 ValSerArgProCysSerCysThrHis 125
DB 932 ATCTCTCGTCAATAGAAAACATCCAC 958

RESULT 15
US-10-296-115-693
; Sequence 693, Application US/10296115
; Publication No. US20040053248A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq Inc
; TITLE OF INVENTION: No. US20040053248A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 784PCT
; CURRENT APPLICATION NUMBER: US/10/296,115
; PRIOR FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: US09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/552,317
; PRIOR FILING DATE: 2000-04-25

; NUMBER OF SEQ ID NOS: 1478
; SEQ ID NO 693
; LENGTH: 1671
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-296-115-693

Alignment Scores:
Pred. No.: 0.109 Length: 1671
Score: 90.00 Matches: 37
Percent Similarity: 38.00% Conservativeness: 20
Best Local Similarity: 24.67% Mismatches: 62
Query Match: 12.53% Indels: 31
DB: 13 Gaps: 4

US-10-005-907-2 (1-135) x US-10-296-115-693 (1-1671)
QY 5 LeuLeuArgLysLeuSerCysLeuGlnGlnLysLysProLysLysGlyAsnPro 24
DB 889 CTCCTACCAATCCTCTTGTCCCTGGAGCGCAGGAGCAAGCGCCAGACCAAGCAGAG 948
QY 25 AspGluGluArgLysArgGlnGluMetThrThrPheGluArgLysLeuGlnAspGlnAsp 44
DB 949 CAAGGAGTGGACACAGCAGGCGGCGACACAGACACAGAGAGAGAGAGAGAGAGAG 1008
QY 45 LysLysSerGlnGluValSerSerThrSerAsnGlnGlnLysLeuGlnAsnGlySerGlySer 64
DB 1009 CAGGAAGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1068
QY 65 GluGlu-----ValCysTyrThrValIleAsnHisIleProHisGlnArgSer 80
DB 1069 AAGGAGGAGCGGAGCGGTGTCTCAGCTGCAGACAGACTCAGACCCCAAGTTTCACTCT 1128
QY 81 ----SerLeuSerSerAsnAspAspGlyTyr----- 89
DB 1129 GAATCTCTATCTTCTTAACCCCTTCTCTTGTCTCCCGGTACGAGAAAGTAGAGTCTACT 1188
QY 90 -----GluAsnIleAspSerLeuThrArgLysValArgGlnPheArgGlu--- 104
DB 1189 CCTATGATTAATGAGAAACATCCAGAGGTCATTCGATCAGCCCGAGAAATAGATGAAATG 1248
QY 105 -----ArgSerGluThrGluTyrAlaLeuLeuArg 114
DB 1249 AATGAAATATATGATGAGAACTCTCTACTGGAGAAACCAAAACCCCTGGCAGGCTCTCGCAG 1308
QY - 115 -ThrSerValSerArgProCysSerCys 123
DB 1309 CTGCCCCACACAGAGCGCTTGTCTGTGC 1336

Search completed: April 22, 2004, 19:43:59
Job time : 345 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model
Run on: April 22, 2004, 17:08:28 ; Search time 81 Seconds
(without alignments)
924.918 Million cell updates/sec

Title: US-10-005-907-2
Perfect score: 715
Sequence: 1 MGNYLRLKSLCIGENQKKK.....SVSRPCSTHERDYEVVPH 135

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n_model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10005907/runat_21042004_113755_11461/app_query.fasta_1.327
-DB=Issued Patents_NA -QFMI=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPC5=0
-LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10005907 @CIGN 1 1 85 @runat_21042004_113755_11461 -NCPUS=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: /cgn2_6/ptodata/2/ina/5A COMB.seq.*
- 2: /cgn2_6/ptodata/2/ina/5B COMB.seq.*
- 3: /cgn2_6/ptodata/2/ina/6A COMB.seq.*
- 4: /cgn2_6/ptodata/2/ina/6B COMB.seq.*
- 5: /cgn2_6/ptodata/2/ina/PCTUS COMB.seq.*
- 6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	96	13.4	2301	1	US-08-306-691B-23
2	96	13.4	2301	4	Sequence 23, Appli
3	96	13.4	2301	5	Sequence 3, Appli
4	90	12.6	8257	4	PCT-US93-06251-78
5	90	12.6	8503	4	Sequence 78, Appli
6	89.5	12.5	1071	4	US-09-595-684B-30
7	86.5	12.1	1302	4	Sequence 30, Appli
8	85.5	12.0	6755	3	US-09-620-312D-130
9	84	11.7	5394	3	Sequence 118, App
10	84	11.7	8146	4	US-08-630-915A-17
11	84	11.7	11917	4	Sequence 17, Appli
12	82.5	11.5	598	4	US-08-931-999-4
					Sequence 4, Appli
					Sequence 1, Appli
					Sequence 725, App
					Sequence 32, Appli
					Sequence 271, App

13	82.5	11.5	6008	1	US-07-789-315A-5	Sequence 5, Appli
14	82.5	11.5	6008	1	US-08-005-002C-5	Sequence 5, Appli
15	82.5	11.5	6008	1	US-08-487-203A-5	Sequence 5, Appli
16	82	11.5	4262	4	US-09-521-511C-10	Sequence 10, Appli
17	81.5	11.4	563	4	US-09-621-976-3759	Sequence 2759, Ap
18	81	11.3	2885	4	US-10-164-595-39	Sequence 39, Appli
19	80	11.2	1428	2	US-08-786-999-3	Sequence 2, Appli
20	80	11.2	1428	2	US-08-989-987-8	Sequence 8, Appli
21	80	11.2	1428	4	US-09-216-387-2	Sequence 2, Appli
22	80	11.2	1494	1	US-08-585-758A-3	Sequence 3, Appli
23	80	11.2	1494	1	US-08-977-818-3	Sequence 3, Appli
24	80	11.2	1494	2	US-08-670-274B-3	Sequence 3, Appli
25	80	11.2	1494	3	US-09-146-187-3	Sequence 3, Appli
26	80	11.2	1494	4	US-09-023-655-997	Sequence 997, App
27	80	11.2	5761	4	US-09-620-312D-147	Sequence 147, App
28	80	11.2	5857	4	US-09-220-132-79	Sequence 79, Appli
29	80	11.2	6107	4	US-09-976-594-557	Sequence 557, App
30	79.5	11.1	512	4	US-09-621-976-3471	Sequence 3471, Ap
31	79.5	11.1	843	2	US-08-933-750C-67	Sequence 67, Appli
32	79.5	11.1	843	3	US-09-234-613-67	Sequence 67, Appli
33	79.5	11.1	4226	4	US-09-620-312D-480	Sequence 480, App
34	78.5	11.1	6002	4	US-09-345-882-4	Sequence 4, Appli
35	78.5	11.0	1491	3	US-08-999-774A-7	Sequence 7, Appli
36	78	10.9	1056	4	US-09-134-001C-1550	Sequence 1550, Ap
37	78	10.9	2139	3	US-09-026-343-1	Sequence 1, Appli
38	78	10.9	2139	4	US-09-362-871-1	Sequence 1, Appli
39	78	10.9	2791	4	US-09-570-367C-1	Sequence 1, Appli
40	78	10.9	2791	4	US-09-915-524-1	Sequence 1, Appli
41	77.5	10.8	2421	4	US-09-417-137-52	Sequence 52, Appli
42	77	10.8	2128	4	US-09-620-312D-197	Sequence 197, App
43	77	10.8	2623	4	US-09-540-824-23	Sequence 23, Appli
44	76.5	10.7	1238	4	US-09-183-861-75	Sequence 75, Appli
45	76.5	10.7	1238	4	US-09-022-765-75	Sequence 75, Appli

ALIGNMENTS

RESULT 1

US-08-306-691B-23
; Sequence 23, Application US/08306691B
; Patent No. 5734039
; GENERAL INFORMATION:
; APPLICANT: Calabretta, Bruno
; APPLICANT: SKORSKI, Tomasz
; TITLE OF INVENTION: ANTISENSE
; TITLE OF INVENTION: OLIGONUCLEOTIDES TARGETING COOPERATING ONCOGENES
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seidel, Gonda, Lavorigna & Monaco, P.C.
; STREET: Two Penn Center, Suite 1800
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/306.691B
; FILING DATE: September 15, 1994
; CLASSIFICATION: 514
; PRIOR APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 8321-8
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549

TELEX: No. 5734039e
 INFORMATION FOR SEQ ID NO: 23:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2301 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear

US-08-306-691B-23

Alignment Scores:

Pred. No.: 0.0132 Length: 2301
 Score: 96.00 Matches: 29
 Percent Similarity: 41.41% Conservativeness: 24
 Best Local Similarity: 22.66% Mismatches: 51
 Query Match: 13.43% Indels: 24
 DB: 1 Gaps: 3

US-10-005-907-2 (1-135) x US-08-306-691B-23 (1-2301)

14 GluAenGlnLysLysProLysGlyAsnProAspGluArgLysArgGlnGluMet 33
 17 GAGGAGCAGGAGGAGCAGGAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGCAG 76
 34 ThrThrPheGluArgLysLysGlnAspGlnAspLysLysSerGlnGluValSerThr 53
 77 GAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 136
 54 SerAenGlnGluAenGlnGlySerGlySerGlnGluValCys----- 68
 137 GAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 196
 69 -----TyrThrValIleAsn 73
 197 AGGAACCGGAGCGGAGCAGTAGCTGGTGGGCACCATGCTGGGATCACCACCATCGAG 256
 74 HistileProHisGlnArgSerSerLeuSerSerAsn---AspAspGlyTyrGluAsnIle 92
 257 GCGGTGAAGCGCAGATCCAGGTCGTGAGCAGGAGGAGGAGGAGGAGGAGGAGGAG 316
 93 AspSerLeuThrArgLysVal-----ArgGlnPheArgGluArgSerGluThrGlu 109
 317 GAGCGCTCCAGCAGAGAGTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 376
 110 TyrAlaLeuLeuArgThrSerVal 117
 377 GTGGCTCTTGACCGTAGGATC 400

RESULT 2

US-09-167-206-3
 Sequence 3, Application US/09167206A
 Patent No. 6476193
 GENERAL INFORMATION:
 APPLICANT: Nandabalan, Krishnan
 APPLICANT: Schulz, Vincent P.
 APPLICANT: Yang, Meijia
 TITLE OF INVENTION: N1K1 PROTEIN AND N1K1 PROTEIN COMPLEXES
 FILE REFERENCE: 15966-521 N1K1 protein complexes
 CURRENT APPLICATION NUMBER: US/09/167,206A
 CURRENT FILING DATE: 1998-10-06
 NUMBER OF SEQ ID NOS: 26
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 3
 LENGTH: 2301
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (233)...(2155)

Alignment Scores:

Pred. No.: 0.0132 Length: 2301
 Score: 96.00 Matches: 29

Percent Similarity: 41.41% Conservativeness: 24
 Best Local Similarity: 22.66% Mismatches: 51
 Query Match: 13.43% Indels: 24
 DB: 4 Gaps: 3

US-10-005-907-2 (1-135) x US-09-167-206-3 (1-2301)

14 GluAenGlnLysLysProLysGlyAsnProAspGluArgLysArgGlnGluMet 33
 17 GAGGAGCAGGAGGAGCAGGAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGCAG 76
 34 ThrThrPheGluArgLysLysGlnAspGlnAspLysLysSerGlnGluValSerThr 53
 77 GAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 136
 54 SerAenGlnGluAenGlnGlySerGlySerGlnGluValCys----- 68
 137 GAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 196
 69 -----TyrThrValIleAsn 73
 197 AGGAACCGGAGCGGAGCAGTAGCTGGTGGGCACCATGCTGGGATCACCACCATCGAG 256
 74 HistileProHisGlnArgSerSerLeuSerSerAsn---AspAspGlyTyrGluAsnIle 92
 257 GCGGTGAAGCGCAGATCCAGGTCGTGAGCAGGAGGAGGAGGAGGAGGAGGAGGAG 316
 93 AspSerLeuThrArgLysVal-----ArgGlnPheArgGluArgSerGluThrGlu 109
 317 GAGCGCTCCAGCAGAGAGTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 376
 110 TyrAlaLeuLeuArgThrSerVal 117
 377 GTGGCTCTTGACCGTAGGATC 400

RESULT 3

PCT-US93-06251-78
 Sequence 78, Application PC/TUS9306251
 GENERAL INFORMATION:
 APPLICANT: Wickstrom, Eric and Rife, Jason P.
 TITLE OF INVENTION: Trivalent Synthesis of Oligonucleotides Containing
 Specific Invention: Stereospecific Alkylphosphonates and Arylphosphonates
 NUMBER OF SEQUENCES: 93
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
 STREET: 400 Garden City Plaza
 CITY: Garden City
 STATE: NY
 COUNTRY: USA
 ZIP: 11530
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US93/06251
 FILING DATE: 19930630
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Digiglio, Frank S.
 REGISTRATION NUMBER: 31,346
 REFERENCE/DOCKET NUMBER: 8586
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 516-742-4343
 TELEFAX: 516-742-4366
 TELEX: 230 901 SANS UR
 INFORMATION FOR SEQ ID NO: 78:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2301 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear

```
MOLECULE TYPE: DNA (genomic)
PCT-US93-06251-78

Alignment Scores:
Pred. No.: 0.0132 Length: 2301
Score: 96.00 Matches: 29
Percent Similarity: 41.41% Conservative: 24
Best Local Similarity: 22.66% Mismatches: 51
Query Match: 13.43% Indels: 24
DB: 5 Gaps: 3

US-10-005-907-2 (1-135) x PCT-US93-06251-78 (1-2301)
QY 14 GluAsnGlnLysLysProLysLysGlyAsnProAspGluGluArgLysArgGlnGluMet 33
Db 17 GAGGAGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 76
QY 34 ThrThrPheGluArgLysLeuGlnAspGlnAspLysLysSerGlnGluValSerThr 53
Db 77 GAGGAGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 136
QY 54 SerAsnGlnGluAsnGlnGlySerGlySerGluGluValCys 68
Db 137 GAGGAGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 196
QY 69 -----TyrThrValIleAsn 73
Db 197 AGNAACCGGAGCGGAGCAGTAGTGGTGGGCGACCATGGCTGGGATCACCACCTCGAG 256
QY 74 HistleProHisGlnArgSerSerLeuSerSerAsn---AspAspGlyTyrGluAsnIle 92
Db 257 GCGGTGAAGCGCAGATCCAGGTTCCTCAGCAGCAGGAGCAGATGTCAGAGGAGCGAGCT 316
QY 93 AspSerLeuThrArgLysVal-----ArgGlnPheArgGluArgSerGluThrGlu 109
Db 317 GAGCGCTCCAGCGAGAGTTGAGGAGAGAAAGCGCGCGGAGCAGGAGGAGGAGGAG 376
QY 110 TyrAlaLeuLeuArgThrSerVal 117
Db 377 GTGGCTCTTGAACCGTAGGATC 400

RESULT 4
US-09-595-684B-30
; Sequence 30, Application US/09595684B
; Patent No. 6544766
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Ohashi, Cara
; APPLICANT: Sakowicz, Roman
; APPLICANT: Vaisberg, Eugeni
; APPLICANT: Wood, Kenneth
; APPLICANT: Yu, Ming
; TITLE OF INVENTION: Human kinesins and methods of producing
; TITLE OF INVENTION: and purifying human kinesins
; FILE REFERENCE: cytop036
; CURRENT APPLICATION NUMBER: US/09/595,684B
; PRIOR FILING DATE: 2002-06-24
; PRIOR APPLICATION NUMBER: 09/295,612
; PRIOR FILING DATE: 2000-04-20
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 8257
; TYPE: DNA
; ORGANISM: Human
US-09-595-684B-30

Alignment Scores:
Pred. No.: 0.551 Length: 8257
Score: 90.00 Matches: 25
Percent Similarity: 49.57% Conservative: 32
Best Local Similarity: 21.74% Mismatches: 44
Query Match: 12.59% Indels: 14

DB: 4 Gaps: 4
US-10-005-907-2 (1-135) x US-09-595-684B-30 (1-8257)
QY 5 LeuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLysProLysLysGlyAsnPro 24
Db 3433 CTTTCTAGGACCTGTGACAGCTGGCAGAGTTGAAAGAAAATAAAGGAAAGAGCCAG 3492
QY 25 AspGluGluArgLysArgGlnGluMetThrThrPheGluArgLysLeuGlnAspGlnAsp 44
Db 3493 CAACTCCAAAGAAAACAGCACTTCTTAATGTACAAGAGAGATGAGTGATGCGAG 3552
QY 45 LysLysSerGlnGluValSerSerThrSerAsnGln---GluAsnGluAsnGlySerGly 63
Db 3553 AAAAAGATTATGAATAGAGATTTAAAGATGAATTAAGAAACAAGAA----- 3603
QY 64 SerGluGluValCysTyrThrValIleAsnHisIleProHisGlnArgSerSerLeuSer 83
Db 3604 -----TTGACATTGGAAACATATGGAACACAGAGAGGCTTGAGTTGGCT 3645
QY 84 SerAsn---AspAspGlyTyrGluAsnIleAspSerLeuThrArgLys-----Val 99
Db 3646 CAGAACTTAATGAATAATTATGAGGAAGTGAATCTATACCAAGAAAGAAAGTTCTA 3705
QY 100 ArgGlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArg 114
Db 3706 AAGGAATTACAGAGTCTATTGAAACACAGAGAGACCCCTTAGA 3750

RESULT 5
US-09-620-312D-130
; Sequence 130, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenchua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunding
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 130
; LENGTH: 8503
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (91)..(8082)
US-09-620-312D-130

Alignment Scores:
Pred. No.: 0.576 Length: 8503
Score: 90.00 Matches: 25
```

Percent Similarity: 49.57% Conservative: 32
 Best Local Similarity: 21.74% Mismatches: 44
 Query Match: 12.59% Indels: 14
 DB: 4 Gaps: 4

US-10-005-907-2 (1-135) x US-09-620-312D-130 (1-8503)

QY 5 LeuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysProLysGlyAsnPro 24
 DB 3433 CTTTCTAGGACCTGTGACAGACTGGCAGAGTTGAAGAAAACCTAAGAAAGAGAGCCAG 3492
 QY 25 AspGluGluArgLysArgGlnGluMetThrThrPheGluArgLysLeuGlnAspGluAsp 44
 DB 3493 CAACCTCAAGAAAACACCAACAACCTTCTTAATGTACAAGAGAGATGAGTGAGATCCAG 3552
 QY 45 LysLysSerGlnGluValSerSerThrSerAsnGln---GluAsnGluAsnGlySerGly 63
 DB 3553 AAAAGATTATGAATATAGAGAAATTAAGATGAAATTAAGAACAAAGAA----- 3603
 QY 64 SerGluGluValCysTyThrValIleAsnHisIleProHisGlnArgSerSerLeuSer 83
 DB 3604 -----TTGACATTGGAACATATGGAACACAGAGAGCGCTTGAGTTGGCT 3645
 QY 84 SerAsn---AspAspGlyTyGluAsnIleAspSerLeuThrArgLys-----Val 99
 DB 3646 CAGAAACTTAATGAAATATGAGAGAGTGAATCTATAACCAAGAAAGAAAAGTTCTA 3705
 QY 100 ArgGlnPheArgGluArgSerGluThrGluTyzAlaLeuLeuArg 114
 DB 3706 AAGGAATTACAGAGAGTCTTTGAAACACAGAGAGACCACCTTAGA 3750

RESULT 6
 US-09-205-258-118
 ; Sequence 118, Application US/09205258
 ; Patent No. 6525174
 ; GENERAL INFORMATION:
 ; APPLICANT: Young et al.
 ; TITLE OF INVENTION: 207 Human Secreted Proteins
 ; FILE REFERENCE: PZ007P1
 ; CURRENT APPLICATION NUMBER: US/09/205,258
 ; CURRENT FILING DATE: 1998-12-04
 ; EARLIER APPLICATION NUMBER: PCT/US98/11422
 ; EARLIER FILING DATE: 1998-06-04
 ; EARLIER APPLICATION NUMBER: 60/048,885
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/049,375
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,881
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,880
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,896
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/049,020
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,876
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,895
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,884
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,894
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,971
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,964
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,882
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,899
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,893
 ; EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,900
 EARLIER FILING DATE: 1997-06-06
 EARLIER APPLICATION NUMBER: 60/048,901
 EARLIER FILING DATE: 1997-06-06
 EARLIER APPLICATION NUMBER: 60/048,892
 EARLIER FILING DATE: 1997-06-06
 EARLIER APPLICATION NUMBER: 60/048,915
 EARLIER FILING DATE: 1997-06-06
 EARLIER APPLICATION NUMBER: 60/049,019
 EARLIER FILING DATE: 1997-06-06
 EARLIER APPLICATION NUMBER: 60/048,970
 EARLIER FILING DATE: 1997-06-06
 EARLIER APPLICATION NUMBER: 60/048,972
 EARLIER FILING DATE: 1997-06-06
 EARLIER APPLICATION NUMBER: 60/048,916
 EARLIER FILING DATE: 1997-06-06
 EARLIER APPLICATION NUMBER: 60/049,373
 EARLIER FILING DATE: 1997-06-06
 EARLIER APPLICATION NUMBER: 60/048,875
 EARLIER FILING DATE: 1997-06-06
 EARLIER APPLICATION NUMBER: 60/049,374
 EARLIER FILING DATE: 1997-06-06
 EARLIER APPLICATION NUMBER: 60/048,917
 EARLIER FILING DATE: 1997-06-06
 EARLIER APPLICATION NUMBER: 60/048,949
 EARLIER FILING DATE: 1997-06-06
 EARLIER APPLICATION NUMBER: 60/048,974
 EARLIER FILING DATE: 1997-06-06
 EARLIER APPLICATION NUMBER: 60/048,883
 EARLIER FILING DATE: 1997-06-06
 EARLIER APPLICATION NUMBER: 60/048,897
 EARLIER FILING DATE: 1997-06-06
 EARLIER APPLICATION NUMBER: 60/048,898
 EARLIER FILING DATE: 1997-06-06
 EARLIER APPLICATION NUMBER: 60/048,962
 EARLIER FILING DATE: 1997-06-06
 EARLIER APPLICATION NUMBER: 60/048,963
 EARLIER FILING DATE: 1997-06-06
 EARLIER APPLICATION NUMBER: 60/048,877
 EARLIER FILING DATE: 1997-06-06
 EARLIER APPLICATION NUMBER: 60/048,878
 EARLIER FILING DATE: 1997-06-06
 EARLIER APPLICATION NUMBER: 60/070,923
 EARLIER FILING DATE: 1997-12-18
 EARLIER APPLICATION NUMBER: 60/092,921
 EARLIER FILING DATE: 1998-07-15
 EARLIER APPLICATION NUMBER: 60/094,657
 EARLIER FILING DATE: 1998-07-30
 NUMBER OF SEQ ID NOS: 1227
 ; SOFTWARE: Patent in Ver. 2.0
 ; SEQ ID NO 118
 ; LENGTH: 1071
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-205-258-118

Alignment Scores:
 Pred. No.: 0.0303 Length: 1071
 Score: 89.50 Matches: 41
 Percent Similarity: 36.26% Conservative: 21
 Best Local Similarity: 23.98% Mismatches: 56
 Query Match: 12.52% Indels: 53
 DB: 4 Gaps: 7

US-10-005-907-2 (1-135) x US-09-205-258-118 (1-1071)

QY 10 SerCysLeuGlyGluAsnGlnLysProLysGlyAsnProAspGluArgLys 29
 DB 47 TCCATCTCTGAGCTCCATGGAGAGCCACCAGCTCGGTGACCAGGAGACTCGGCGCAAG 106
 QY 30 ArgGlnGluMetThrThrPheGluArgLysLeuGlnAspGlnAspLysSerGlnGlu 49
 DB 107 GCCCGAGACAGCGCCGCCCTGGAAGAACTACAAGACAGAGAAACAA---CAGAAA 163

```
QY 50 ValSerSerThrSerAsnGlnGluValPheProHis 135
Db 164 GTGAGATTTCGTAAAGAGATGGAGAGAGGTGTTCAGATTTCATTCAGACAGTGGCGAG 223
QY 65 GluGluValCysTyrThrValIleAsnHisTleProHisGlnArgSer 80
Db 224 ATCAAGAAAAGATTTCAGCCCATGACAGATC-----GAGAGGACATACATCATAT 277
QY 81 -----SerLeuSerSerAsnAspAsp----- 87
Db 278 GTGGTGAAGTGGTGGCTGCATCATCTCTCTCTTGGGGAAGATGATGACTGTGCGTAT 337
QY 88 -----GlyTyrGluAsnIleAspSerLeu----- 95
Db 338 GTCATGATCTTCAAAAAGGAGTTTGCACCTCAGATGAAGAGCTGACTCTTACCGTCGT 397
QY 96 -----ThrArgLysValArgGlnPheArgGlu 104
Db 398 GGAGAGGAATGGACCCCGACAGAGCTGAGGAGAGCGGAAGCTGAAGGAGCTGGCCAG 457
QY 105 ArgSerGluThrGluTyrAlaIleuLeuArgThrSerValSerArgProCysSerCysThr 124
Db 458 AGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 511
QY 125 HisGluHisAspTyrGluValValPheProHis 135
Db 512 -----GACTACAGGACAGTACAGCCAC 535

RESULT 7
US-08-630-915A-17
; Sequence 17, Application US/08630915A
; Patent No. 6309820
; GENERAL INFORMATION:
; APPLICANT: SPARKS, Andrew B.
; APPLICANT: HOFFMAN, No. 6309820h
; APPLICANT: KAY, Brian K.
; APPLICANT: FOWLES, Dana M.
; APPLICANT: MCCONNELL, Stephen J.
; TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
; TITLE OF INVENTION: DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
; TITLE OF INVENTION: USING SAME
; NUMBER OF SEQUENCES: 227
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/630,915A
; FILING DATE: 03-APR-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Mierock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-174
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1302 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
```

```
; MOLECULE TYPE: DNA
US-08-630-915A-17
Alignment Scores:
Pred. No.: 0.101 Length: 1302
Score: 86.50 Matches: 34
Percent Similarity: 41.67% Conservativity: 21
Best Local Similarity: 12.76% Mismatches: 50
Query Match: 12.10% Indels: 27
DB: 4 Gaps: 4

US-10-005-907-2 (1-135) x US-08-630-915A-17 (1-1302)
QY 14 GluAsnGlnLysLysProLysLysGlyAsnProAspGluGluArgGlnGluMet 33
Db 562 GAGAACCGCGCTGAGGAGAGCGGCTGCCGAAGAGAGGAGCGGCGTTGGAGGAG 621
QY 34 ThrThrPheGluArgLysLeuGlnAspGlnAspLysLysSerGlnGluVal----- 50
Db 622 GAGCGACGAGAGCGGAGGCTGCAGGAGGCTGCCACGTCGAGCAGCGGTACCAGGAACAG 681
QY 51 -----SerSerThrSerAsnGlnGluAsnGluAsnGlySerGly 63
Db 682 CACAGATCAGCTGGAGCCCGGAGCAGGACAGTGTGAGCCAGCAGGAGCCGTTTCAAGG 741
QY 64 SerGluGluValCysTyrThrValIleAsnHisTleProHisGlnArg----- 79
Db 742 ACCAGACAGGAGTGGGAGTCTGCTGGGACAGCAGGCGCCACACCCACGAGAGATTTCAAG 801
QY 80 -----SerSerLeuSerSerAsnAspAspGlyTyrGluAsn 91
Db 802 CAGAAGAAAGGCAATGTCCACCACCTCTGTCCAGCTCCGAGCCGCGCAAGCTGAGG 861
QY 92 IleAspSerLeuThrArgLysValArgGlnPheArgGluArgSerGluThrGluTyrAla 111
Db 862 AGCCCTTCTCTGCAAGCAACTCACTCAA-----CCAGAAACCTCTCTACGCG 909
QY 112 LeuLeuArgThrSer---ValSerArgProCysSer 122
Db 910 CGAGAGCCACAGCTCTCTCTCCCGGCTGACGCA 945

RESULT 8
US-08-931-999-4
; Sequence 4, Application US/08931999
; Patent No. 6043219
; GENERAL INFORMATION:
; APPLICANT: Iandolo, John J.
; APPLICANT: Crupper, Scott S.
; TITLE OF INVENTION: Broad Spectrum Chemotherapeutic Peptide
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hovey, Williams, Timmons & Collins
; STREET: 2405 Grand Boulevard, Suite 400
; CITY: Kansas City
; STATE: Missouri
; COUNTRY: U.S.A.
; ZIP: 64108
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/931,999
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/710,561
; FILING DATE: 19-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Collins, John M.
; REGISTRATION NUMBER: 26,262
; REFERENCE/DOCKET NUMBER: 25043-A
```


ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte ID No. 6673549 4295277CB1
NAME/KEY: unsure
LOCATION: 8083
OTHER INFORMATION: a, t, c, g, or other
US-09-976-594-725

Alignment Scores:
Pred. No.: 3.35 Length: 8146
Score: 84.00 Matches: 32
Percent Similarity: 38.02% Conservative: 14
Best Local Similarity: 26.45% Mismatches: 47
Query Match: 11.75% Indels: 28
DB: 4 Gaps: 5

US-10-005-907-2 (1-135) x US-09-976-594-725 (1-8146)
QY 13 GlycLluAsnGlnLysLysPro-----LysLysGlyAsnProAsp 25
DB 2886 GGGGATGTGAACAGACACCTTGAGGAGCTGGAGCCCGCTCGAGAGCAGGAGGTAGACGAC 2945
QY 26 GluGluArgLysArgGlnGluMetThrPheGluArgLysLeuGlnAspGlnAspLys 45
DB 2946 ATTGAATAATTGAAGATGAAGAGCGCGTGTTCAGAGATCTTCAGAGACTGGAGAC--- 3002
QY 46 LysSerGlnGluValSerSerThrSerAsnGlnGluAsnGlySerGlySerGlu 65
DB 3003 ---TATGAGAGAGAGCGGAACTGAGGAGGCTGAGGAGCCAGAGAGGATGGGAGGAA 3059
QY 66 GluValCysTyrThrValIleAsnHisIlePro----- 76
DB 3060 CACGTATGTGTGAGCGGCTCCAGACACAGCCCACTGAGGATGAGGAAAGTGCAGAGCG 3119
QY 77 -----HisGlnArgSerSerLeuSerSerAsnAspAsp---Gly 88
DB 3120 GAGGCTGATGATCATCATCAGGAGAGAGGAGTCTGTGCCAGTGGGATGACCGAGCC 3179
QY 89 TyrGluAsnIleAspSerLeuThrArgLysValArgGlnPheArgGluArgSerGluThr 108
DB 3180 GAAGAGACATGATGAGGAGGAGGAGGAGTCTGTGCCAGTGGGATGACCGAGCC 3233
QY 109 Glu 109
DB 3234 GAG 3236

RESULT 11
US-09-566-921-32
; Sequence 32, Application US/09566921
; Patent No. 682888
; GENERAL INFORMATION:
; APPLICANT: Loring, Jeanne F.
; APPLICANT: Tingley, Debora W.
; APPLICANT: Edwards, Carla M.
; TITLE OF INVENTION: GENES EXPRESSED IN ALZHEIMER'S DISEASE
; FILE REFERENCE: PA-0024 US
; CURRENT APPLICATION NUMBER: US/09/566,921
; CURRENT FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PERL Program
; SEQ ID NO 32
; LENGTH: 11917
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6682888 475473.1
; NAME/KEY: unsure
; LOCATION: 9384, 10290-10321, 10325
; OTHER INFORMATION: a, t, c, g, or other
US-09-566-921-32

Alignment Scores:
Pred. No.: 5.92 Length: 11917
Score: 84.00 Matches: 32
Percent Similarity: 38.02% Conservative: 14
Best Local Similarity: 26.45% Mismatches: 47
Query Match: 11.75% Indels: 28
DB: 4 Gaps: 5

US-10-005-907-2 (1-135) x US-09-566-921-32 (1-11917)
QY 13 GlycLluAsnGlnLysLysPro-----LysLysGlyAsnProAsp 25
DB 2887 GGGGATGTGAACAGACACCTTGAGGAGCTGGAGCCCGCTCGAGAGCAGGAGGTAGACGAC 2946
QY 26 GluGluArgLysArgGlnGluMetThrPheGluArgLysLeuGlnAspGlnAspLys 45
DB 2947 ATTGAATAATTGAAGATGAAGAGCGCGTGTTCAGAGATCTTCAGAGACTGGAGAC--- 3003
QY 46 LysSerGlnGluValSerSerThrSerAsnGlnGluAsnGlySerGlySerGlu 65
DB 3004 ---TATGAGAGAGAGGAGGAGAACTGAGGAGGCTGAGGAGCCAGAGAGGATGGGAGGAA 3060
QY 66 GluValCysTyrThrValIleAsnHisIlePro----- 76
DB 3061 CACGTATGTGTGAGCGGCTCCAGACACAGCCCACTGAGGATGAGGAAAGTGCAGAGCG 3120
QY 77 -----HisGlnArgSerSerLeuSerSerAsnAspAsp---Gly 88
DB 3121 GAGGCTGATGATCATCATCAGGAGAGAGGAGTCTGTGCCAGTGGGATGACCGAGCC 3180
QY 89 TyrGluAsnIleAspSerLeuThrArgLysValArgGlnPheArgGluArgSerGluThr 108
DB 3181 GAAGAGACATGATGAGGAGGAGGAGGAGTCTGTGCCAGTGGGATGACCGAGAG 3234
QY 109 Glu 109
DB 3235 GAG 3237

RESULT 12
US-08-961-527-271/c
; Sequence 271, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 271:

SEQUENCE CHARACTERISTICS:
LENGTH: 598 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-961-527-271

Alignment Scores:

Pred. No.: 0.107 Length: 598
Score: 82.50 Matches: 51
Percent Similarity: 41.71% Conservative: 27
Best Local Similarity: 27.27% Mismatches: 37
Query Match: 11.54% Indels: 73
DB: 4 Gaps: 11

US-10-005-907-2 (1-135) x US-08-961-527-271 (1-598)

QY 3 AsnTyr-----LeuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLys 18
DB 534 AACTACAGATTAGTACAGCTTACGAGGAGGAGAAAGCTGATTGCCAGACCTCAAGAAA 475
QY 19 ProLysLysGlyAsnProAspGluArgLysArgGlnGluMetThrThrPheGluArg 38
DB 474 CCGAARAAG-----TCCAGAAG-----TCATCGAGTTCGCAG 442
QY 39 LysLeuGlnAspGlnAspLysLysSerGlnGluValSer----- 51
DB 441 AAATCTCGACCAAGAGAGAAAATCAAGAGAAAATCGATGATCATCGTCTATCTAGTTG 382
QY 52 -----SerThrSer-AsnGlnGluAsnGlnGly----- 61
DB 381 AGATTGGCAAGATTACGGCTTCTATATCGAGATCGAGAGTCTAGCAGAAACCGCA 322
QY 62 -----SerGlySerGluGluValCysTyrThrVa 71
DB 321 GTCAGACCTACGACGATTGATTCGCACTATCATAGTTCTTA-AACCTCTTGCACCAACAT 263
QY 71 lileasnHisleProHisGlnArgSerSerLeuSerSerAsnAspAspGlyTyrGluAs 91
DB 262 TTTATCCAAATTCCTCAATCAGAAATCATCATATCG----- 226
QY 91 nileaspSerLeuThrArgLysValArgGlnPhe-----Ar 103
DB 225 -ATTCCATCTCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 167
QY 103 gLuarGSerGluThrGluTyrAlaLeuLeu----- 113
DB 166 GAAAGCTTCACTACTCTTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 107
QY 114 ----ArgThrSerValSerArgProCysSer-----CysThrHisGlu---HisAspTy 129
DB 106 TGAAAAAACAAAGTATTTCACGA---TGCAGCTACCGGTGTTCTTATCGGCTCATCTCTCT 50
QY 129 rGluValValPheProHis 135
DB 49 CTATCTCTTTTCACTCAT 31

RESULT 13

US-07-789-915A-5

Sequence 5, Application US/07789915A

Patent No. 5212058

GENERAL INFORMATION:

APPLICANT: Baker, Rohan T.
APPLICANT: Tobias, John W.
APPLICANT: Varshavsky, Alexander
TITLE OF INVENTION: Ubiquitin-Specific Proteases
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESSES:
ADDRESS: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Militia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: U.S.A.

RESULT 14

US-08-005-002C-5

Sequence 5, Application US/08005002C

Patent No. 5494818

GENERAL INFORMATION:

APPLICANT: Baker, Rohan T.
APPLICANT: Tobias, John W.

ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/07789,915A
APPLICATION NUMBER: US/07789,915A
FILING DATE: 19911108
CLASSIFICATION: 435
NAME: Brook, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: MIT-5091AA
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 6008 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 983..4774
US-07-789-915A-5

Alignment Scores:

Pred. No.: 3.36 Length: 6008
Score: 82.50 Matches: 32
Percent Similarity: 36.80% Conservative: 14
Best Local Similarity: 25.60% Mismatches: 60
Query Match: 11.54% Indels: 20
DB: 1 Gaps: 3

US-10-005-907-2 (1-135) x US-07-789-915A-5 (1-6008)

QY 6 LeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLysProLysLysGlyAsnProAsp 25
DB 5112 ATGAGAAACAAAGCCGCTCTCAA-GATGATAATGTCAAGGAAGTCAATGGAATGATGAT 5170
QY 26 GluGluArgLysArgGlnGluMetThrThrPheGluArgLysLeu----- 40
DB 5171 ACTAAAGAGAGGAACCAAG 5230
QY 41 -----GlnAspGlnAspLysLysSerGlnGluValSerSerThrSerAsnGlnGlu 57
DB 5231 GATGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5290
QY 58 AsnGluAsnGlySerGlySerGluGluValCysTyrThrValIleasnHisleProHis 77
DB 5291 GATGAG 5332
QY 78 GlnArgSerSerLeuSerSerAsnAspAspGlyTyrGluAsnIleAspSerLeuThrArg 97
DB 5333 CATAGAGAAAGTCTCGGTTAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5392
QY 98 Lys-----ValArgGlnPheArgGluArgGluThrGluTyrAlaLeu 112
DB 5393 AAAATTTATGTTGAGGATGCTAATAACTTTTCTGATTTGATGAGAGACTACTCACACTTTT 5452
QY 113 LeuArgThrSerVal 117
DB 5453 AAAGAGAAATGTA 5467

APPLICANT: Varshavsky, Alexander
TITLE OF INVENTION: Ubiquitin-Specific Proteases
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS: ADDRESSEE: Kevin M. Farrell, P.C.
STREET: P.O. Box 999
CITY: York Harbor
STATE: Maine
COUNTRY: U.S.A.

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

Alignment Scores:	3.36	Length:	6008
Pred. No.:	82.50	Matches:	32
Score:	35.80%	Conservative:	10
Percent Similarity:	2.60%	Mismatches:	60
Best Local Similarity:	2.80%	Indels:	20
Query Match:	11.54%	Gaps:	3
DR:	1		

US-10-005-907-2 (1-135) x US-08-005-002C-5 (1-6008)

Qy		6	LeuArgIysLeuSerCysLeuGluGlnValSerProLysLysGlyAsnProAsp	25
Db		5112	ATGAGGAACAACGCCGTCTCAA-GATGATATGTCAGGAAAGTCAATGGAAATGATGAT	5170
Qy		26	GluGluArgLysAsgGlnGluMetThrThrPheGluArgLysLeu-----	40
Db		5171	ACTAAAGAAGAGGACACAGNAGAACGCACAGCTAGATGATTATTTTGGAGATGACAA	5230
Qy		41	-----GlnAspGlnAspIylsSerGlnGluValSerSerThrSerAsnGlnGlu	57
Db		5231	GATGACGATGATGATGATGATGTTATAAAAATCGGAGACTGAAAAAGATGATAGTAGT	5290
Qy		58	AsnGlnAsnGlySerGlySerGluGluValCysTyThrValIleAsrHisIleProHis	77
Db		5291	GATGAAGACGACGGGAGAGAA-----ATCACCATAGAAAGTCGT	5332
Qy		78	GlnArgSerSerLeuSerSerAsnAspGlyTyrgluAsnIleAspSerLeuThrArg	97
Db		5333	CATGAGAAGTGCTCGGGTTGATGATGATGACGAGGAGGACCCATGTCACCCGA	5392
Qy		98	Lys-----ValargGlnPheArgGluArgSerGluThrgluTyAlaLeu	112
Db		5393	AAATTTTATGTTGAGCATGCTAAATAACCTTTTCTGATCTTGATGAGACTACTACACACTTTT	5452

```

113 LeuArgThrSerVal 117
::|||
Db 5453 AAAGAGGAAATGTA 5467

RESULT 15
US-08-487-203A-5
; Sequence 5, Application US/08487203A
; Parent No. 5683904
; GENERAL INFORMATION:
; APPLICANT: Baker, Rohan T.
; APPLICANT: Tobias, John W.
; APPLICANT: Varshavsky, Alexander
; TITLE OF INVENTION: Ubiquitin-Specific Proteases
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kevin M. Farrell, P.C.
; STREET: P.O. Box 999
; CITY: York Harbor
; STATE: Maine
; COUNTRY: U.S.A.
; ZIP: 03911
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,203A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/005,002
; FILING DATE: 15-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Farrell, Kevin M.
; REGISTRATION NUMBER: 35,505
; REFERENCE/DOCKET NUMBER: MIT-5091A3Z
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 207-363-0558
; TELEFAX: 207-363-0528
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6008 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 983..4774
US-08-487-203A-5

```

Alignment Scores:	3.36	Length:	6008
Pred. No.:	82.50	Matches:	32
Score:	36.80%	Conservative:	14
Percent Similarity:	25.60%	Mismatches:	60
Best Local Similarity:	11.54%	Indels:	20
Query Match:	1	Gaps:	3
DB:			

US-10-005-907-2 (1-135) X US-08-487-203A-5 (1-6008)

QY		6	LeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLysProLysLysGlyAsnProasp	25
			:::: ::::	
Dd		5112	ATGAGGAACAACGGCGTCTCAA-GATGTAATGTCAAGGAAGTCATGGAATGATCAT	5170
			:::: ::::	
QY		26	GluGluArgLysArsgInGluMetThrPheGluArgLysLeu-----	40
			:::: ::::	
Dd		5171	ACTAAAGAAGAGGACCAAGAGAAGACGCACATAGATGATTATTTCGAGATGACAAT	5230
			:::: ::::	
QY		41	-----GlnAspGlnAspLysLysSerGlnGluValSerSerThrSerAsnGlnGlu	S7
			:::: ::::	
Dd		5231	GATGACGATGATGATGATGATGTTTAAAAAATCGGAGACTGAAAATAAGTGATGATAGT	5290
			:::: ::::	

```
Qy 58 AsnGluAsnGlySerGlySerGluGluValCysTyrThrValIleAsnHisIleProHis 77
Db 5291 GATGAGACGACGAGGAGAGAAAT-----ATCACCATAGAGATCGT 5332
Qy 78 GlnArgSerLeuSerSerAsnAspGlyTyrGluAsnIleAspSerLeuThrArg 97
Db 5333 CATAGAGAACTCTCGGGTTAGATGATGATGACAGAGGAGCAAGCCATGACACCCGA 5392
Qy 98 Lys-----ValArgGlnPheArgGluArgSerGluThrGluTyrAlaLeu 112
Db 5393 AAATTTTATGGTGAGGATGCTATAACTTTTCTGATCTTGATGAGACTACTCACACTTTT 5452
Qy 113 LeuArgThrSerVal 117
Db 5453 AAAGAGGAAATGTA 5467
```

Search completed: April 22, 2004, 18:40:34
Job time : 99 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 22, 2004, 11:22:40 ; Search time 351 Seconds
(without alignments)
4901.767 Million cell updates/sec

Title: US-10-005-907-1_COPY_25_429

Perfect score: 405
Sequence: 1 atgggaattatctctctgcg.....atgaagtgtgtttccacac 405

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_29Jan04:*

1: Geneseqn1980s:*

2: Geneseqn1990s:*

3: Geneseqn2000s:*

4: Geneseqn2001as:*

5: Geneseqn2001bs:*

6: Geneseqn2002s:*

7: Geneseqn2003as:*

8: Geneseqn2003bs:*

9: Geneseqn2003cs:*

10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	405	100.0	667	9	ADC30760 Human nov
2	405	100.0	3762	6	ABN81319 Human mas
3	58	14.3	732	7	ACC59951 Human IRA
4	47.2	11.7	1643	6	ABL99791 Human sec
5	46.6	11.5	775	7	ABX05108 Human nov
6	41.6	10.3	6446	4	AS46328 Tumour su
7	41.6	10.3	18988	4	AA54341 Tumour su
8	41.6	10.3	18988	6	ABL32700 Human imm
9	41.6	10.3	18988	6	ABL34508 Human met
10	41.6	10.3	18988	6	ABL70203 Chemocall
11	41	10.1	1145	4	AAK17837 Human bra
12	41	10.1	1996	4	AAK05119 Human bra
13	41	10.1	11924	9	ADC30201 Human nov
14	41	10.1	11951	9	ADC30200 Human nov
15	41	10.1	14994	7	ADL19440 CDNA erco
16	40.4	10.0	6523	6	ABN80121 Human che
17	40.4	10.0	6523	9	ABE84216 Human lym
18	40.4	10.0	6523	9	ABE84140 Human lym
19	40.2	9.9	4186	2	AZ411320 Human nor
20	40	9.9	2921	3	AAZ77829 Human can
21	40	9.9	3402	4	AAH54452 S. epider
22	40	9.9	3668	4	AAH54402 S. epider
23	40	9.9	4100	4	AAH54798 S. epider

ALIGNMENTS

RESULT 1

ADC30760

ID ADC30760 standard; CDNA; 667 BP.

XX

AC ADC30760;

XX

DT 18-DEC-2003 (first entry)

XX

DE Human novel cDNA sequence, SEQ ID NO:842.

XX

KW Human; diagnostic; drug screening; forensics; gene mapping;

KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;

KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;

KW ulcers; osteoporosis; autoimmune disease; cancer;

KW molecular weight marker; food supplement; antiparkinsonian; nootropic;

KW neuroprotective; anti-anemic; anticoagulant; thrombolytic; vulnary;

KW antiulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;

KW gene therapy; chromosome 1; gene; ss.

XX

OS Homo sapiens.

XX

PN WO2003029271-A2.

XX

PD 10-APR-2003.

XX

PF 24-SEP-2002; 2002WO-US030474.

XX

PR 24-SEP-2001; 2001US-0324631P.

XX

PA (HYSE-) HYSEQ INC.

XX

PI Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;

PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;

PI Haley-Vicente D, Drmanac RT;

XX

XX WPI; 2003-371981/35.

DR P-PSDB; ADC31731.

XX

PT New polynucleotide and polypeptide useful for diagnosing, preventing or

PT treating conditions such as neurodegenerative diseases, anemias, platelet

PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or

PT cancer.

XX

PS Claim 1; SEQ ID NO 842; 1185pp; English.

XX

CC The invention relates to 971 novel human cDNA sequences (ADC29919-

CC ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The
 CC invention also relates to nucleic acid sequences over 99% identical with
 CC the novel human cDNAs. The invention additionally encompasses expression
 CC vectors and host cells comprising a nucleic acid of the invention; the
 CC recombinant production of a polypeptide of the invention; an antibody
 CC against a polypeptide of the invention; a method of detecting
 CC polynucleotides or polypeptides of the invention; and methods of
 CC identifying a compound which binds to a polypeptide of the invention. The
 CC invention further discloses methods of preventing, treating or
 CC ameliorating a medical condition; kits comprising polynucleotide probes
 CC and/or monoclonal antibodies for carrying out the methods of the
 CC invention; methods for the identification of compounds that modulate the
 CC expression or activity of the polynucleotide and/or polypeptide; and 767
 CC contig sequences corresponding to the cDNA sequences of the invention
 CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
 CC -ADC33394). The nucleic acids and polypeptides of the invention are
 CC useful in diagnostics, drug screening, forensics, gene mapping, in the
 CC identification of mutations responsible for genetic disorders or other
 CC traits, for assessing biodiversity, and in producing many other types of
 CC data and products dependent on DNA and amino acid sequences. They are
 CC also used for treating diseases such as Parkinson's disease, Alzheimer's
 CC disease and other neurodegenerative diseases, anaemia, platelet
 CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
 CC cancer. The nucleic acids may also be used as hybridisation probes or
 CC primers, and in the recombinant production of a protein. The polypeptides
 CC are also useful in generating antibodies, as molecular weight markers,
 CC and as food supplements. The present sequence represents a specifically
 CC claimed human cDNA sequence of the invention. Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 667 BP; 216 A; 147 C; 154 G; 150 T; 0 U; 0 Other;

Query Match 100.0%; Score 405; DB 9; Length 667;
 Best Local Similarity 100.0%; Pred. No. 2.9e-107;
 Matches 405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGAATTAATCTCTCGGAAACTCAGTTGCTCGGAGAGATCAAAAGAGCCCAAG 60
 DB 186 ATGGGAATTAATCTCTCGGAAACTCAGTTGCTCGGAGAGATCAAAAGAGCCCAAG 245

QY 61 AAAGGAACCCAGATGAGGAAGAAACCGGAGGAATGACTACATTTGAAAGAACTT 120
 DB 246 AAAGGAACCCAGATGAGGAAGAAACCGGAGGAATGACTACATTTGAAAGAACTT 305

QY 121 CAAGATCAGATACAAAGCCAGAGATTTTCACCTCTTATCAATCAGGAAGAGAAAT 180
 DB 306 CAAGATCAGATACAAAGCCAGAGATTTTCACCTCTTATCAATCAGGAAGAGAAAT 365

QY 181 GGCAGTGGTTCTGAGAGAGTGTGCTACACTGTCTATTAATCAATCCATCCATCAGATCC 240
 DB 366 GGCAGTGGTTCTGAGAGAGTGTGCTACACTGTCTATTAATCAATCCATCCATCAGATCC 425

QY 241 TCCCTGAGCTCCATGATGAGTGGTATGAGACATTCATCCCTCAGAGAGAGTGA 300
 DB 426 TCCCTGAGCTCCATGATGAGTGGTATGAGACATTCATCCCTCAGAGAGAGTGA 485

QY 301 CAGTTTAGAGAAAGTCCAGACACAGAAATATGCCCTTCTTAGGACTTCTGTAGTGGCCT 360
 DB 486 CAGTTTAGAGAAAGTCCAGACACAGAAATATGCCCTTCTTAGGACTTCTGTAGTGGCCT 545

QY 361 TGTTCCTGACCCATGACATGATATGAAAGTGTGTTCCACAC 405
 DB 546 TGTTCCTGACCCATGACATGATATGAAAGTGTGTTCCACAC 590

RESULT 2
 ABN81319
 ID ABN81319 standard; cDNA; 3762 BP.

XX ABN81319;
 XX

DT 30-AUG-2002 (first entry)
 XX Human mast cell related gene MC1 SEQ ID NO 1.
 DE
 XX Human; mast cell; MC; antiallergic; antiinflammatory; antiasthmatic;
 KW vasotropic; dermatological; allergy; hypersensitivity; rhinitis; asthma;
 KW gene; ss.
 XX Homo sapiens.
 OS
 XX Key Location/Qualifiers
 FT 25..432
 CDS /*tag= a
 /product= "MC1"
 XX WO200246389-A2.
 PN 13-JUN-2002.
 PD 07-DEC-2001; 2001WO-US046180.
 XX 08-DEC-2000; 2000US-0251835P.
 PR 14-MAR-2001; 2001US-0275479P.
 PR 28-MAR-2001; 2001US-0279115P.
 PR 02-APR-2001; 2001US-0280143P.
 XX (UNIO) UCB SA.
 PA Nocka K, Pirozzi G, Einstein R;
 FI WPI; 2002-508560/54.
 XX P-PSDB; ABB77569.
 DR
 XX Novel isolated nucleic acids that are differentially expressed in mast
 PT cells in patients with allergic hypersensitivity, encoding proteins
 PT associated with mast cell regranulation and allergic hypersensitivity.
 XX
 PS Claim 1; Page 95-97; 119pp; English.

XX The invention relates to isolated nucleic acid (ABN81319-ABN81324),
 CC corresponding to genes differentially expressed in mast cells following
 CC activation or in patients with allergic hypersensitivity disease. (I)
 CC that encodes proteins (ABB77569-ABB77575) (II) or a protein fragment of
 CC (II) if at least 6 amino acids. (III) is useful for identifying binding
 CC partners. (I) or (II) is useful for diagnosing or treating a disease
 CC state (e.g. allergic hypersensitivity, seasonal rhinitis, asthma,
 CC urticaria or atopic dermatitis or mastocytosis) in a subject which
 CC involves determining the level of expression of (I) or (II). A computer
 CC system, comprising a database containing information identifying the
 CC expression level in a tissue or at least one mast cell of (I), is useful
 CC for preventing information to identify the relative expression level of
 CC (I). (II) is used as a marker to detect, diagnose or identify an allergic
 CC response in a patient. The protein can also serve as a target that
 CC modulate gene expression or activity and as an antigen to raise
 CC polyclonal or monoclonal antibodies. (II) is useful for identifying
 CC agents that modulate expression of the protein or agents, such as
 CC agonists or antagonists. The agonists or antagonists are useful for
 CC modulating biological activity and function of (II) and thus are useful
 CC for alleviating disease conditions such as allergic hypersensitivity,
 CC seasonal rhinitis, asthma, urticaria, atopic dermatitis or mastocytosis
 XX
 SQ Sequence 3762 BP; 1220 A; 672 C; 680 G; 1190 T; 0 U; 0 Other;

Query Match 100.0%; Score 405; DB 6; Length 3762;
 Best Local Similarity 100.0%; Pred. No. 5.4e-107;
 Matches 405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGAATTAATCTCTCGGAAACTCAGTTGCTCGGAGAGATCAAAAGAGCCCAAG 60
 DB 25 ATGGGAATTAATCTCTCGGAAACTCAGTTGCTCGGAGAGATCAAAAGAGCCCAAG 84

QY 61 AAAGGAACCCAGATGAGGAAGAAACCGGAGGAATGACTACATTTGAAAGAACTT 120

85	AAAGGAAACCCGATGAGGAAAGAAACGGCAGGAAATGACTACATTTGGAAGAAAACCTT	144
121	CAAGATCAAGATATAGAAAGAAGCCAGAGTTTATCCCTCTTAAATCAGGAAACGAGAAAT	180
145	CAAGATCAAGATATAGAAAGAAGCCAGAGTTTATCCCACTTTCTAATCAGGAAACGAGAAAT	204
181	GSCAGTGGTCTTGGAAGAAGTGTGTACACTGTCTAATTAATCACATATCCCCCATCAGAGATCC	240
205	GGCAATGGTCTTGGAAGAAGTGTGTCAACTGTCTAATTAATCAATCCTCCCATCAGAGATCC	264
241	TCCCTTGAGCTCCAAATGATGATGGCTATGAGAAACATTTGACTCCCTTCACAAGGAAAGTGAGA	300
265	TCCCTTGAGCTCCAAATGATGATGGCTATGAGAAACATTTGACTCCCTTCACAAGGAAAGTGAGA	324
301	CAGTTTTAGAGAAAGTCCAGAGACAGAAATATGCCCTCTTTAGGACTTCTGTTTAGTAGGCCT	360
325	CAGTTTTAGAGAAAGTCCAGAGACAGAAATATGCCCTCTTTAGGACTTCTGTTTAGTAGGCCT	384
361	TGTTCTGCACCCATGAGCATGATTAATGAAGTTGTGTTTCCACAC	405
385	TGTTCTGCACCCATGAGCATGATTAATGAAGTTGTGTTTCCACAC	429

RESULT 3	
ACC59951	
ID	ACC59951 standard; cDNA; 732 BP.
XX	
XX	ACC59951;
XX	
DT	07-JUL-2003 (first entry)
XX	
XX	Human IRAP-28 encoding cDNA SEQ ID NO:63.
XX	
XX	Human; immune response associated protein; IRAP; anti-HIV; antiallergic
XX	antiinflammatory; antianaemic; antiparkinsonian; nootropic; antithyroid
XX	anticonvulsant; antiarteriosclerotic; antiasthmatic; immunosuppressive;
XX	cytostatic; hepatotropic; dermatological; antidiabetic; nephrotropic;
XX	antigout; chymotryptic; neuroprotective; osteopathic; aniaethetic;
XX	antiparasitic; antihelminthic; antipsoriatic; trophantic; protozoacide;
XX	antirheumatic; haemostatic; antibacterial; virucide; ophthalmologic;
XX	fungicide; gene therapy; immune system disorder; neurological disorder;
XX	developmental disorder; muscle disorder; cell proliferative disorder;
XX	infection; gene; ss.

AD	Homo sapiens.
CS	
XX	Location/Qualifiers
XX	Key
PH	134..670
FT	/tag= a
FT	/product= "IRAP-28"
FT	
FT	
XX	
FN	WC2003025542-A2.
XX	
PD	27-MAR-2003.
XX	
PF	19-SEP-2002; 2002WO-US029979.
XX	
PR	21-SEP-2001; 200IUS-0324034P.
PR	05-OCT-2001; 200IUS-0327395P.
PR	12-OCT-2001; 200IUS-0328223P.
PR	19-OCT-2001; 200IUS-0342810P.
PR	09-NOV-2001; 200IUS-0344468P.
PR	21-NOV-2001; 200IUS-0332140P.
PR	07-DEC-2001; 200IUS-0340282P.
PR	09-JAN-2002; 2002US-0347693P.
PR	20-FEB-2002; 2002US-0358279P.
PR	01-MAR-2002; 2002US-0361088P.
PR	15-MAR-2002; 2002US-0364494P.
PR	10-MAY-2002; 2002US-0379876P.
PR	11-JUN-2002; 2002US-0388180P.
XX	(INCY-) INCYTE GENOMICS INC.
PA	
XX	

Ho A, Baughn ME, Becha SD, Burford N, Elliott VS, Emerling BM;
 Forsythe IJ, Corvrad AE, Griffin JA, Hafalia AJA, Honchell CD,
 Ison CH, Burrill JD, Blake JJ, Lal PG, Lee EA, Marquis JP,
 Lehr-Mason PM, Lee S, Sprague WW, Swarnakar A, Tang YJ, Tran B;
 Tran UK, Bhatia U, Chawla NK, Warren B, Zheng W, Xu Y, Yue H;
 WPI; 2003-363161/34.
 P-PSDB; ABR43232.
 New human immune response associated proteins and polynucleotides, useful
 for diagnosing, treating or preventing immune system disorders, e.g. AIDS
 or anemia, cell proliferative disorders, e.g. cancer, or neurological
 disorders.
 Claim 5; Page 208; 213pp; English.
 ACS59924 to ACS59958 encode the human immune response associated proteins
 given in ABR43205 to ABR43239, designated IRAP-1 to IRAP-29 (I). (I) have
 anti-HIV, anti-allergic, anti-inflammatory, antianemic, anti-parkinsonian,
 neutropic, anticonvulsant, antiarteriosclerotic, antiasthmatic, antigout,
 immunosuppressive, antithyroid, cytostatic, hepatotropic, dermatological,
 antididiabetic, nephrotropic, thymometric, neuroprotective, osteopathic,
 antihistatic, antiparasitic, antihelminthic, antiproliferative, uropathic,
 ophthalmological, antirheumatic, haemostatic, antibacterial, varicidal,
 protozoacidal and fungicide activities, and can be used in gene therapy.
 Human IRAP polynucleotides, agonists and antagonists are useful for
 diagnosing, treating or preventing disorders associated with aberrant
 expression of IRAP, such as immune system disorders (e.g. AIDS, asthma,
 allergies, autoimmune thyroiditis, contact dermatitis, Crohn's disease,
 diabetes mellitus, glomerulonephritis, Good pasture's syndrome, gout,
 Graves' disease, Hashimoto's thyroiditis, irritable bowel syndrome,
 multiple sclerosis, osteoarthritis, osteoporosis, pancreatitis, Reiter's
 syndrome, rheumatoid arthritis, Sjogren's syndrome, Parkinson's disease or
 neurological disorders (e.g. Alzheimer's disease, renal tubular acidosis, myocardiitis
 or mental retardation), muscle disorder (e.g. cardiomyopathy, myocardiitis
 or dermatomyositis), cell proliferative disorders (e.g. arteriosclerosis,
 atherosclerosis, cirrhosis, hepatitis, paroxysmal nocturnal
 haemoglobinuria, polycythemia vera, psoriasis, primary
 thrombocytopaenia or cancer), or viral, bacterial, fungal, parasitic,
 protozoan or helminthic infections
 Sequence 732 BP; 213 A; 178 C; 178 G; 163 T; 0 U; 0 Other;

Query Match	14.3%;	Score 58;	DB 7;	Length 732;	
Best Local Similarity	54.4%;	Pred. No. 1.3e-06;			
Matches 162;	Conservative 0;	Mismatches 130;	Indels 6;	Gaps 2;	

QY	114	AAAACCTTCAAGATCAAGATAGAAGAAAGCGCAAGAGTTTTCATCCCACTTCTTAATCAGGAAA	173
DB	298	AAAGAGGCAAGATTTCCCAAAACGAAATGAAGAAATGTCTACTCTCCATCCAGGACAA	357
QY	174	CGAAGATGGCAGTGGTTCTGAAGAAGTGTGCTACACTGTCTATTAAATACA---TCCGCCCA	230
DB	358	TGTTGACCCAGACCTPACTCAGAGGAGCTGTGCTATACCCCTCATCAATCATCGGGTTCCTGTG	417
QY	231	TCAGAGATCTCCCTGTGAGCTCCATGATGATGGCTATGAGAACATTGACTTCCTCCCTCACAAG	290
DB	418	TACAAGGCCATCAGGGAACTCTGCTGAAGAGTACTATGAGAATGTTCCTCGAAAGCTGA	477
QY	291	GAAGTGTAGACAGATGTTTAGAGAAAGGTTCAGAGACAGAAATATGCCCTTCT---TAGGACTTC	347
DB	478	GAGACCCAGAGATGCTTGGAGGAACTGAGACTGAGTATTTCATCTTACATATAGCCTTC	537
QY	348	TGTTAGTAGGCGCTTGTCTCTGCAACCCATGAGCATGATTATGAAGTTGTGTTCCACAC	405
DB	538	TACAGACCCAGGCGATGCCCGATCCCCAGAGATGAATATGAACCTTCTCATGCGCTCAC	595

RESULT 4	
ABL979791	
ID	ABL979791 standard; cDNA; 1643 BP.
XX	

AC ABL99791;

DT 03-OCT-2002 (first entry)

DE Human secretory polynucleotide (septm) 46.

XX Human; ss; gene; secretory protein; secretory polynucleotides; SPTM;
XX SPTM-related disease; somatic gene therapy; germline gene therapy;
XX severe combined immunodeficiency; intracellular parasite protection;
XX fungal parasite; protozoan parasite; cell proliferative disorder; cancer;
XX immune disorder; AIDS; neurological disorder; Parkinson's disease;
XX motor neuron disorder; demyelinating disease; multiple sclerosis;
XX meningitis; abscess; prion diseases; cerebral palsy;
XX neuroskeletal disorder; peripheral nervous system disorder;
XX dermatomyositis; polymyositis; myopathy; myasthenia gravis;
XX mental disorder; Tourette's syndrome.

XX Homo sapiens.

XX WO200220756-A2.

XX 14-MAR-2002.

XX 30-AUG-2001; 2001WO-US027297.

XX 05-SEP-2000; 2000US-0229747P.

XX 05-SEP-2000; 2000US-0229748P.

XX 05-SEP-2000; 2000US-0229749P.

XX 05-SEP-2000; 2000US-0229750P.

XX 05-SEP-2000; 2000US-0229751P.

XX 05-SEP-2000; 2000US-0230016P.

XX 05-SEP-2000; 2000US-0230583P.

XX 06-SEP-2000; 2000US-0230505P.

XX 06-SEP-2000; 2000US-0230514P.

XX 06-SEP-2000; 2000US-0230515P.

XX 06-SEP-2000; 2000US-0230517P.

XX 06-SEP-2000; 2000US-0230518P.

XX 06-SEP-2000; 2000US-0230519P.

XX 06-SEP-2000; 2000US-0230596P.

XX 06-SEP-2000; 2000US-0230597P.

XX 06-SEP-2000; 2000US-0230599P.

XX 06-SEP-2000; 2000US-0230610P.

XX 06-SEP-2000; 2000US-0230864P.

XX 06-SEP-2000; 2000US-0230865P.

XX 06-SEP-2000; 2000US-0230988P.

XX 06-SEP-2000; 2000US-0230989P.

XX 06-SEP-2000; 2000US-0230990P.

XX 07-SEP-2000; 2000US-0230896P.

XX 07-SEP-2000; 2000US-0230897P.

XX 07-SEP-2000; 2000US-0230951P.

XX 07-SEP-2000; 2000US-0231163P.

XX 07-SEP-2000; 2000US-0231832P.

XX (INCY-) INCYTE GENOMICS INC.

XX Stuart J, Lincoln SE, Altus CM, Dufour GP, Chalup MS, Hillman JL;
XX Jones AL, Yu JY, Wright RJ, Gietzen D, Liu TF, Yap PE, Dahl CR;
XX Momiyama MG, Bradley DL, Rohatgi SD, Harris B, Roseberry AM;
XX Gerstein EH, Feratta CH, David MA, Panzer SR, Flores V, Daffo A;
XX Marwaha R, Chen AJ, Chang SC, Au AP, Inman RR;
XX WPI; 2002-315658/35.
XX P-PSDB; ABB97794.
XX Polynucleotide sequences encoding human secretory proteins useful for
XX gene therapy of e.g. genetic deficiency disorders, cancers, and diseases
XX caused by intracellular parasites.
XX Claim 1; Page 275-276; 585pp; English.
XX The invention comprises the amino acid and coding sequences of human
XX secretory (SPTM) proteins. The SPTM DNA and amino acid sequences are

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

useful for treating a disease or condition associated with the expression
of functional SPTM. The SPTM DNA sequences are useful for somatic or
germline gene therapy to correct a genetic deficiency (e.g. severe
combined immunodeficiency). The SPTM DNA sequences are also useful in
providing protection against intracellular parasites (e.g. fungal
parasites and protozoan parasites). The SPTM DNA and protein sequences
are also useful for diagnosing cell proliferative disorders, cancer, s
immune disorders (e.g. AIDS), neurological disorders (e.g. Parkinson's
disease), motor neuron disorders, demyelinating diseases (e.g. multiple
sclerosis), meningitis, abscesses, prion diseases, cerebral palsy,
neuroskeletal disorders, peripheral nervous system disorders,
dermatomyositis and polymyositis, myopathy, myasthenia gravis, and mental
disorders (e.g. Tourette's syndrome). cDNA sequences ABL99746 - ABL99929
CC represent human secretory polynucleotides of the invention
XX
SQ Sequence 1643 BP; 413 A; 382 C; 384 G; 460 T; 0 U; 4 Other;

Query Match 11.7%; Score 47.2; DB 6; Length 1643;

Best Local Similarity 50.8%; Pred. No. 0.0024;

Matches 165; Conservative 0; Mismatches 15; Indels 6; Gaps 2;

QY 87 ACGCAGGAATGACTACATTGAAAGAAACTTCAAGATCAAGATAAGAAAGCCAGA 146

Db 1236 ATGGAAAAAATACTCAATTTTGAAGAGGCAAGATCCCAAAACGAAAAATGAAAGAT 1295

QY 147 AGTTTCATCCACTTCTAATCAGGAAACGAGATGGCAGTGGTCTGGAAGTGTGCTA 206

Db 1296 GTCATCTACTCCATCCAGCAGGACCAATGTTGACAGACCTACTCAGAGGAGCTGTGCTA 1355

QY 207 CACTGTCTAATATCA---CATCCCCCATCAGAGATCCCTCCCTGAGCTCAATGATGATGG 263

Db 1356 TACCCTCATCAATCATCGGGTTCTGTACAGGCGCATCAGGGAACCTCTGCTGAAGAGTA 1415

QY 264 CTATGAGAACATGACTCCCTCACAAGGAAAGTGAGAGAGTTAGAGAAAGTTCAGAGAC 323

Db 1416 CTATGAGAAATGTTCCCTGCAAGAGTGAAGACCCAGAGAGTCTTGGGAGGAACTGAGAC 1475

QY 324 AGAATATGCCCTTCTTAGGACTTCTGTTAGTAGAGGCTTTG---TTCCTGACCCATGAGCA 380

Db 1476 TGAGTATTCACITTCATANATATGCTTCTACAGACCCAGGCATGCCGATCCCAAGAGA 1535

QY 381 TGATTAAGAGTTGTGTTTCCACAC 405

Db 1536 TGAATTAAGAACTTTTCATGCTTCAC 1560

RESULT 5

ID ABL99746 standard; cDNA; 775 BP.

XX ABL99746;

XX 17-JAN-2003 (first entry)

XX Human novel polynucleotide #123.

XX Human; gene; ss; genetic disorder; gene mapping; medical imaging; cancer;
XX neurodegenerative disorder; lymphoid cell disorder; osteoporosis;
XX Parkinson's disease; Alzheimer's disease; bone degenerative disorder;
XX osteoarthritis; periodontal disease; liver fibrosis; viral infection;
XX fungal infection; bacterial infection; autoimmune disease; diabetes;
XX atopic dermatitis.

XX Homo sapiens.

XX WO200274961-A1.

XX 26-SEP-2002.

XX 14-MAR-2002; 2002WO-US005109.

XX 15-MAR-2001; 2001US-00810173.

XX

(HYSE-) HYSEQ INC.

Tang YT, Zhou P, Goodrich R, Asundi V, Zhang J, Zhao QA, Ren F; Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M; Wehrman T, Wang J, Wang D, Drmanac RT; WPI: 2003-040556/03. P-PSDB; ABU00030.

New isolated polypeptides and polynucleotides, useful for preventing, treating or ameliorating medical conditions, such as cancer, neurodegenerative disorders, lymphoid cell disorders, bone degenerative disorders, and infections.

Claim 1; SEQ ID NO 123; 235pp; English.

The invention relates to human polynucleotides and the polypeptides they encode. The polynucleotides and polypeptides are useful in diagnostics, forensics, gene mapping, medical imaging, identification of mutations, responsible for genetic disorders or other traits, assessing biodiversity and producing many other types of data and products dependent on DNA and amino acid sequences. They are also useful for preventing, treating or ameliorating medical conditions, such as cancer, neurodegenerative disorders (e.g. Parkinson's disease, Alzheimer's disease), lymphoid cell disorders, osteoporosis, osteoarthritis, bone degenerative disorders, periodontal disease, liver fibrosis, infections (e.g. viral, fungal or bacterial) or autoimmune diseases (e.g. diabetes, atopic dermatitis). Sequences ABX04986-ABX05511 represent human polynucleotides of the invention. Note: The sequence data for this patent is not represented in the printed specification but is based on sequence information supplied by the European Patent Office

Sequence 775 BP; 226 A; 202 C; 160 G; 187 T; 0 U; 0 Other;

Query Match 11.5%; Score 46.6; DB 7; Length 775;
Best Local Similarity 50.8%; Pred. No. 0.0027;
Matches 165; Conservative 0; Mismatches 154; Indels 6; Gaps 2

QY 87 ACGGAGGAAATGACTACATTGTGAAGAAACCTTCAAGATCAAGATAAGAAAGCCAGA 146
DB 325 ATGGAAGAAATACTCATTTTTTGAAGAGGCAAGATTCCAAAACGAAATGAAAGAAT 384
QY 147 AGTTTCATCCACTTCTAATCTAGGAACGAGAAATGGCAGTGGTTCCTGAAGAGTGTGCTA 206
DB 385 GTCATCTACTCCCATCCAGCAGGACAATGTTGACCAACCTACTCAGAGGAGCTGTGCTA 444
QY 207 CACTGTCTAATTAATCA--CATCCCCCATCAGAGATCCTCCCTGAGCTCCAATGATGAGG 263
DB 445 TACCCTCATCAATCATCGGGTTCTGTGACAAGGCCATCAGGAACTCTGCTGAAGAGTA 504
QY 264 CTATGAGACATGTACTCCCTCACAAGAAAGTGGACACTTTAGAGAAAGGTCAGAGAC 323
DB 505 CTAATGAGAAATGTTCCCTGCAAGCTGGAGAGACCACAGAGAGTCTTTGGAGGAATCTGAGAC 564
QY 324 AGAATATGCCCTTCTTAGACTTCTGTTAGTAGGCCCTTG---TTCCTGCAACCCATGAGCA 380
DB 565 TGAGTATTCACTTCTACATATGCGCTTCTACAGACCCAGGAGTCCCGCATCCCAAGA 624
QY 381 TGAATTGAAAGTTGTTTCCACAC 405
DB 625 TGAATATGAAGTCTCTCATGCCCTCAC 649

RESULT 6
AAS46328/c
ID AAS46328 standard; DNA; 6446 BP.
XX
AC AAS46328;
XX
XX 18-DEC-2001 (first entry)
XX
XX Tumour suppressor gene derived chemically modified sequence #50.
XX

Human; tumour suppressor gene; oncogene; antitumour; cytostratic; cancer; tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP; cytosine methylation; ds.	Human sapiens.	WO200168912-A2.	20-SEP-2001.	15-MAR-2001; 2001WO-EP002955.	15-MAR-2000; 2000DE-01013847.	06-APR-2000; 2000DE-01013058.	07-APR-2000; 2000DE-01019173.	30-JUN-2000; 2000DE-01032529.	01-SEP-2000; 2000DE-01043826.	(EPIG-) EPIGENOMICS AG.	Olek A, Piepenbrock C, Berlin K;	WPI; 2001-602752/68.	Fragments of chemically modified genes associated with tumor suppressor genes and oncogenes, useful in designing primers and probes for analyzing diseases associated with cytosine methylation state e.g. cancer.	Claim 1; SEQ ID NO 50; 27pp; English.	The invention relates to a nucleic acid comprising a sequence of 18 bases, of a segment of chemically pretreated DNA (Cp DNA) e.g. with bisulphite, of genes associated with tumour suppression and oncogenes having a sequence taken from 536 (actually 533 since numbers 408, 458 and 500 are missing from the sequence listing) sequences (Ss) and sequences complementary to (Gs). The nucleic acid may be a peptide nucleic acid-oligonucleotide (PNA) of at least 9 nucleotides and may form part of a set of probes for detecting the cytosine methylation state and/or single nucleotide polymorphisms and also to be used in an array for analysing diseases associated with CpG dinucleotides e.g. cancers and tumours. The probes can also be used in a method for ascertaining genetic and/or epigenetic parameters for the diagnosis and/or therapy of existing diseases or the predisposition to specific diseases, by analysing cytosine methylations. The parameters may be compared to another set of genetic and/or epigenetic parameters, the differences serving as basis for diagnosis and/or prognosis events which are disadvantageous to patients. The present sequence is one of the 533 genomic sequences derived from tumour suppressor genes and oncogenes. Sequences with even numbered Seq ID numbers are the complementary sequence of the corresponding odd numbered sequence (e.g. ID 2 and ID1, ID 536 and ID 535, except for those whose partner sequence is missing). Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences	Sequence 6446 BP; 1491 A; 208 C; 1786 G; 2951 T; 0 U; 0 Other;	Query Match Best Local Similarity 10.3%; Score 41.6; DB 4; Length 6446; Matches 77; Conservative 56.6%; Pred. No. 0.16; Mismatches 59; Indels 0; Gaps 0;	39 AGAGATCAAAAGAACCCCAAGAAAGAACCCAGTCAGGAAGAAACCGCAGGAAT 98 1454 ACAAAAAAAAAAAAAACCCCAACCAATATAAA CAAAAAAAAAAAAACGCAAAAAA 1395 99 GACTACATTTGAAAGAAACCTTCAAGATCAAGATAAGAAAAAGCCAAAGATTTCATCCAC 158 1394 AACGACGAAAAAATAAAAAATCTTAATTAATAAAAAAACAATAAAAAAAGCA 1395 159 TTCTAATCAGGAAAC 174 1334 TAAAAAACCAGAAAC 1319
--	----------------	-----------------	--------------	-------------------------------	-------------------------------	-------------------------------	-------------------------------	-------------------------------	-------------------------------	-------------------------	----------------------------------	----------------------	--	---------------------------------------	--	--	--	--

XX	SQ	Sequence 6446 BP; 1491 A; 208 C; 1796 G; 2951 T; 0 U; 0 Other;
		Query Match 10.3%; Score 41.6; DB 4; Length 6446;
		Best Local Similarity 56.6%; Pred No. 0.16; Indels 0; Gaps 0;
		Matches 77; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	39	AGAGAATCAAAGAAGCCCAAGAAAGGAACCCAGATGAGGAAGAAGAAAACGGCGAGGAAT 98
Db	1454	ACAATAAAAAAAAAAACCCCAACCATTAATAA CAIAAAAAAAAAAAAAAACCACAAAAAAA 1395
QY	99	GACTACATTTCGAAAGAAAACCTTCAAGATCAAGATAAGAGAAAGCCAAGAAAGTTTCATCCAC 158
Db	1394	AACGACAGAAAATAAAAAAAAAATTTCAATTAAATAAAAAAAAAACAAAAAAAATAAACGA 1335
QY	159	TTCCTAATCAGGA AAC 174
Db	1334	TAAAAAAAACGAAAC 1319

RESULT 7
AAS46341/c
ID AAS46341 standard; DNA; 18988 BP.
AC AAS46341;
XX
XX
DE 18-DEC-2001 (first entry)
XX
XX
DE Tumour suppressor gene derived chemically modified sequence #63.
XX
XX Human; tumour suppressor gene; oncogene; antitumour; cytostatic; cancer;
KW tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;
KW cytosine methylation; ds.
XX
XX Homo sapiens.
OS
XX
XX WO200168912-A2.
PN
XX
XX 20-SEP-2001.
PD
XX
XX 15-MAR-2001; 2001WO-EP002955.
PF
XX
XX 15-MAR-2000; 2000DE-01013847.
PR
XX 06-APR-2000; 2000DE-01013058.
PR
XX 07-APR-2000; 2000DE-01013173.
PR
XX 30-JUN-2000; 2000DE-01032529.
PR
XX 01-SEP-2000; 2000DE-01043826.
PR
XX
XX (EPIG-) EPIGENOMICS AG.
PA
XX
XX Olek A, Piepenbrock C, Berlin K;
PI
XX WPI; 2001-602752/68.
XX
XX Fragments of chemically modified genes associated with tumor suppressor
PT genes and oncogenes, useful in designing primers and probes for analyzing
PT diseases associated with cytosine methylation state e.g. cancer.
PT
XX
XX Claim 1; SEQ ID NO 63; 27pp; English.
PS
XX
XX The invention relates to a nucleic acid comprising a sequence of 18
CC bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with
CC bisulphite, of genes associated with tumour suppression and oncogenes
CC having a sequence taken from 536 (actually 533 since numbers 408, 458 and
CC 500 are missing from the sequence listing) sequences (Ss) and sequences
CC complementary to (Ss). The nucleic acid may be a peptide nucleic acid-
CC oligomer (PNA) of at least 9 nucleotides and may form part of a set of
CC probes for detecting the cytosine methylation state and/or single
CC nucleotide polymorphisms and also to be used in an array for analysing
CC diseases associated with CpG dinucleotides e.g. cancers and tumours. The
CC probes can also be used in a method for ascertaining genetic and/or
CC epigenetic parameters for the diagnosis and/or therapy of existing
CC diseases or the predisposition to specific diseases, by analysing
CC cytosine methylations. The parameters may be compared to another set of
CC genetic and/or epigenetic parameters, the differences serving as basis
CC for diagnosis and/or prognosis events which are disadvantageous to
CC patients. The present sequence is one of the 533 genomic sequences
CC derived from tumour suppressor genes and oncogenes. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 18988 BP; 4826 A; 567 C; 4681 G; 8914 T; 0 U; 0 Other;
SQ

QY 39 AGAGAAATCAAAAAGAAGCCCAAGAAAGGAAACCCCGATGAGGAAGAAAAACGGCAGGAAT 98
DB 18346 AAAATATTAAATAATTATATATATATATATATATCTTTCACCAACCATATAACCCCACAAAAA 19387

Qy	99	GACTACATTGTGAAGAAACACTTCAGATCAAGATGAAGAAAGCCAGAGAGTTTCATCCAC	158
Db	18286	AACAATTTACAANAANAANAACCAAAATTTCAACAACTAATATCAAAACAACATTTTTCTCTC	18227
Qy	159	TTCTAATCAGGAAACGAGAAATGCGAGTGGTTTCTGAAGAAGTGTGCTACACTGTCAATTAA	218
Db	18226	TCTAATCAATTAACAAACAACTTTACCAACTATATCAATATTTTATCACCACCTACTATCAA	18167
Qy	219	TCACATCCCCCATCAGAGATCCTCCCTGAGCTCCCAATGATGATGGCTATGAGACATTGA	278
Db	18166	CAACCAATTAATTCACAAATAATCTCCTATCTTCAAAACATCGAACACCTATCACCCCTAA	18107
Qy	279	CTCCCTCCACAAGGAAAGTGAGACA	302
Db	18106	ACCCCTCCATAAANAATAACACA	18083
RESULT 8			
ABL32700/c			
ID	ABL32700 standard; DNA; 18988 BP.		
XX	ABL32700;		
AC	26-MAR-2002 (first entry)		
DT	Human immune system associated gene SEQ ID NO: 673.		
XX	Human; immune system disease; cytosine methylation; antiasthmatic;		
KW	antiarteriosclerotic; antianaemic; cytosolic; nontropic;		
KW	neuroprotective; anti-HIV; anticonvulsant; ophthalmological;		
KW	anti-rheumatic; antiarthritic; antidiabetic; antipsoriatic;		
KW	anti-inflammatory; cancer; eye disease; arteriosclerosis; anaemia;		
KW	acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;		
KW	neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;		
KW	ds.		
XX	Homo sapiens.		
OS	Homo sapiens.		
XX	W0200200928-A2.		
PN	03-JAN-2002.		
FD	02-JUL-2001; 2001WO-EP007537.		
PF	30-JUN-2000; 2000DE-01032529.		
XX	01-SEP-2000; 2000DE-01043826.		
PR	(EPIG-) EPIGENOMICS AG.		
XX	Olek A, Piepenbrock C, Berlin K;		
FA	WPI; 2002-130909/17.		
XX	Nucleic acid comprising fragment of chemically modified gene, useful for		
PT	diagnosis and treatment of diseases associated with abnormal cytosine		
FT	methylation.		
PT	Claim 1; SEQ ID NO 673; 32pp + Sequence Listing; German.		
XX	The present invention provides a number of human immune system associated		
CC	genes which are modified by the methylation of cytosines. The sequences		
CC	can be used in the diagnosis and treatment of immune system disorders,		
CC	including eye diseases such as retinopathy, neovascular glaucoma and		
CC	macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid		
CC	leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,		
CC	rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel		
CC	diseases. The present sequence is a gene of the invention		
XX	Sequence 18988 BP: 4826 A; 567 C; 4681 G; 8914 T; 0 U; 0 Other;		

Query Match 10.3%; Score 41.6; DB 6; Length 18988;
Best Local Similarity 47.3%; Pred. No. 0.24;
Matches 125; Conservative 0; Mismatches 139; Indels 0; Gaps 0;


```
XX SQ Sequence 1996 BP; 487 A; 410 C; 341 G; 758 T; 0 U; 0 Other;
Query Match 10.1%; Score 41; DB 4; Length 1996;
Best Local Similarity 48.5%; Pred. No. 0.16;
Matches 113; Conservative 0; Mismatches 120; Indels 0; Gaps 0;
QY 62 AAGGAAACCCAGATGAGGAAAGAAACGGCAGGAATGCTACATTTGGAAGAAACTTC 121
DB 1187 ATGAAAGTTCAGAAAGAAACACACACCCCATGAAGTTCTCTGAAACAGCCTAAAGACC 1128
QY 122 AGATCAAGATAGAAAGCCAGAGTTTCATCCACTTCTATCAGGAAGAGGATG 181
DB 1127 AAGAGAAACTCAGAGTTATCTGAAACCTTGGAATTTACTATTTCAGAGAGGAGATCA 1068
QY 182 GCAGTGGTCTGAGGAAGTGTGTACACTGTCTATTAATCAGATCCCCCATCAGAGATCCT 241
DB 1067 AAGAGAGTCAAGAGAAAGAAAGACACATTTTAAAGATAGCCAAAGATATTTCCCT 1008
QY 242 CCTGAGTCCATGATGATGGCTATGAGACATTCAGTCCCTCACAAGGAA 294
DB 1007 CCAGCAGGACCAAGAGAGAGTCTGAGTTTGTGATGACATAGTACTAGTA 955
RESULT 13
ID ADC30201 standard; cDNA; 11924 BP.
XX AC ADC30201;
XX DT 18-DEC-2003 (first entry)
XX DE Human novel cDNA sequence, SEQ ID NO:283.
XX KW Human; diagnostic; drug screening; forensics; gene mapping;
KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;
KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
KW ulcers; osteoporosis; autoimmune disease; cancer;
KW molecular weight marker; food supplement; antiparkinsonian; neurotropic;
KW neuroprotective; antianaemic; anticoagulant; thrombolytic; vulnerary;
KW antilucer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
KW gene therapy; chromosome 3; gene; ss.
XX OS Homo sapiens.
XX PN WO2003029271-A2.
XX PD 10-APR-2003.
XX PF 24-SEP-2002; 2002WO-US030474.
XX PR 24-SEP-2001; 2001US-0324631P.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
PI Haley-Vicente D, Drmanac RT;
XX WPI; 2003-371981/35.
XX DR P-PSDB; ADC31172.
XX PT New polynucleotide and polypeptide useful for diagnosing, preventing or
PT treating conditions such as neurodegenerative diseases, anaemias, platelet
PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
PT cancer.
XX PS Claim 1; SEQ ID NO 283; 1185pp; English.
XX CC The invention relates to 971 novel human cDNA sequences (ADC29919-
CC ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The
CC invention also relates to nucleic acid sequences over 99% identical with
CC the novel human cDNAs. The invention additionally encompasses expression
```

```
CC vectors and host cells comprising a nucleic acid of the invention; the
CC recombinant production of a polypeptide of the invention; an antibody
CC against a polypeptide of the invention; a method of detecting
CC polynucleotides or polypeptides of the invention; and methods of
CC identifying a compound which binds to a polypeptide of the invention. The
CC invention further discloses methods of preventing, treating or
CC ameliorating a medical condition; kits comprising polynucleotide probes
CC and/or monoclonal antibodies for carrying out the methods of the
CC invention; methods for the identification of compounds that modulate the
CC expression or activity of the polynucleotide and/or polypeptide; and 767
CC contig sequences corresponding to the cDNA sequences of the invention
CC (ADC31961-ADC32627) and the polypeptides encoded by the contigs (ADC32628
CC -ADC33394). The nucleic acids and polypeptides of the invention are
CC useful in diagnostics, drug screening, forensics, gene mapping, in the
CC identification of mutations responsible for genetic disorders or other
CC traits, for assessing biodiversity, and in producing many other types of
CC data and products dependent on DNA and amino acid sequences. They are
CC also used for treating diseases such as Parkinson's disease, Alzheimer's
CC disease and other neurodegenerative diseases, anaemia, platelet
CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
CC cancer. The nucleic acids may also be used as hybridisation probes or
CC primers, and in the recombinant production of a protein. The polypeptides
CC are also useful in generating antibodies, as molecular weight markers,
CC and as food supplements. The present sequence represents a specifically
CC claimed human cDNA sequence of the invention. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
```

Sequence 11924 BP; 3959 A; 2706 C; 2427 G; 2832 T; 0 U; 0 Other;

```
Query Match 10.1%; Score 41; DB 9; Length 11924;
Best Local Similarity 48.5%; Pred. No. 0.3;
Matches 113; Conservative 0; Mismatches 120; Indels 0; Gaps 0;
QY 62 AAGGAAACCCAGATGAGGAAAGAAACGGCAGGAATGCTACATTTGGAAGAAACTTC 121
DB 1048 ATGAAAGTTCAGAAAGAAACACACACCCCATGAAGTTCTCTGAAACAGCCTAAAGACC 1107
QY 122 AAGATCAAGATAGAAAGCCAGAGTTTTCATCCACTTCTATCAGGAAGAGGATG 181
DB 1108 AAGAGAAACTCAGAGTTTATCTGAAACCTTGGAAATTAATTTTCAGAGAGGAGATCA 1167
QY 182 GCAGTGGTCTGAGGAAGTGTGTACACTGTCTATTAATCAGATCCCCCATCAGAGATCCT 241
DB 1168 AAGAGAGTCAAGAGAAAGAAAGACACATTTTAAAGATAGCCAAAGATATTTCCCT 1227
QY 242 CCTGAGTCCATGATGATGGCTATGAGACATTCAGTCCCTCACAAGGAA 294
DB 1228 CCAGCAGGACCAAGAGAGAGTCTGAGTTTGTGATGACATAGTACTAGTA 1280
```

RESULT 14

ADC30200

ID ADC30200 standard; cDNA; 11951 BP.

XX AC ADC30200;

XX DT 18-DEC-2003 (first entry)

XX DE Human novel cDNA sequence, SEQ ID NO:282.

XX KW Human; diagnostic; drug screening; forensics; gene mapping;
KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;
KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
KW ulcers; osteoporosis; autoimmune disease; cancer;
KW molecular weight marker; food supplement; antiparkinsonian; neurotropic;
KW neuroprotective; antianaemic; anticoagulant; thrombolytic; vulnerary;
KW antilucer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
KW gene therapy; chromosome 3; gene; ss.
XX OS Homo sapiens.

FN WO2003029271-A2.
XX 10-APR-2003.
XX 24-SEP-2002; 2002WO-US030474.
XX 24-SEP-2001; 2001US-0324631P.
XX (HYSE-) HYSEQ INC.
XX Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
XX Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
XX Haley-Vicente D, Drmanac RT;
XX WPI; 2003-371981/35.
XX P-PSDB; ADC31171.
XX New polynucleotide and polypeptide useful for diagnosing, preventing or
XX treating conditions such as neurodegenerative diseases, anemias, platelet
XX disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
XX cancer.
XX Claim 1; SEQ ID NO 282; 1185pp; English.
XX The invention relates to 971 novel human cDNA sequences (ADC29919-
XX ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The
XX invention also relates to nucleic acid sequences over 99% identical with
XX the novel human cDNAs. The invention additionally encompasses expression
XX vectors and host cells comprising a nucleic acid of the invention; the
XX recombinant production of a polypeptide of the invention; an antibody
XX against a polypeptide of the invention; a method of detecting
XX polynucleotides or polypeptides of the invention; and methods of
XX identifying a compound which binds to a polypeptide of the invention. The
XX invention further discloses methods of preventing, treating or
XX ameliorating a medical condition; kits comprising polynucleotide probes
XX and/or monoclonal antibodies for carrying out the methods of the
XX invention; methods for the identification of compounds that modulate the
XX expression or activity of the polynucleotide and/or polypeptide; and 767
XX contig sequences corresponding to the cDNA sequences of the invention
XX (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
XX -ADC33394). The nucleic acids and polypeptides of the invention are
XX useful in diagnostics, drug screening, forensics, gene mapping, in the
XX identification of mutations responsible for genetic disorders or other
XX traits, for assessing biodiversity, and in producing many other types of
XX data and products dependent on DNA and amino acid sequences. They are
XX also used for treating neurodegenerative diseases, Parkinson's disease, Alzheimer's
XX disease and other neurodegenerative diseases, anaemia, platelet
XX disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
XX cancer. The nucleic acids may also be used as hybridisation probes or
XX primers, and in the recombinant production of a protein. The polypeptides
XX are also useful in generating antibodies, as molecular weight markers,
XX and as food supplements. The present sequence represents a specifically
XX claimed human cDNA sequence of the invention. Note: The sequence data for
XX this patent did not form part of the printed specification, but was
XX obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.
XX Sequence 11951 BP; 3967 A; 2711 C; 2434 G; 2839 T; 0 U; 0 Other;
XX
XX Query Match 10.1%; Score 41; DB 9; Length 11951;
XX Best Local Similarity 48.5%; Pred. No. 0.3;
XX Matches 113; Conservative 0; Mismatches 120; Indels 0; Gaps 0;
XX
XX 62 AAGGAACCCAGATCAGGAAGAAACGGCAGGAAATGACTACATTGGAAGAAACCTC 121
XX 1048 ATGAAGACTCAGAAAGAAACACACACCCATGAAGTTCTCTGACACGCTTAAGACC 1107
XX 122 AAGATCAAGATAAGAAAGCCAGGAAGTTTCATCCATCTTCAATCAGAAACGAGAAATG 181
XX 1108 AAGAGAAACCTCAGAGTTTATCTGAAACCTTGAAGAAATTAATTTTCAAGAGAGAGATCA 1167
XX 182 GCAGTGGTCTTGAGAGAGTGCTGCTACCTGTCATTAATTCATATCCATCCCATCAGAGATCCT 241

Db 1168 AAGAGAGTCAAGAAAGAAAGGAAAGACACTTTTAAAAAGATAGCCACACAGATATTCCTT 1227
Qy 242 CCCTGAGCTCCATGATGATGCTATGAGACATGATGCTCCCTCACAAGGAAA 294
Db 1228 CCAGCAGGACCATATAAGAGAAAGTCTGAGTTTGTGATGACATACTACTAGA 1280
RESULT 15
ACD19440
ID ACD19440 standard; cDNA; 14994 BP.
XX ACD19440;
XX AC
XX DT 25-AUG-2003 (first entry)
XX DE
XX CDNA encoding novel human protein #120.
XX Human; NOV; gene therapy; endocrine related disease; diabetes;
XX metabolism-related disease; obesity; central nervous system disorder;
XX Alzheimer's disease; Parkinson's disease; epilepsy; multiple sclerosis;
XX schizophrenia; depression; autoimmune disorder; inflammatory disorder;
XX psoriasis; allergy; lupus erythematosus; asthma; cancer;
XX inflammatory bowel disease; rheumatoid arthritis; osteoarthritis;
XX colon cancer; lung cancer; liver cancer; breast cancer; ovarian cancer;
XX prostate cancer; brain cancer; melanoma; liver disease; liver cirrhosis;
XX lung disease; emphysema; obstructive pulmonary disease; haemophilia;
XX stroke; infection; Gene; ss.
XX Homo sapiens.
XX OS
XX WO2003023002-A2.
XX PD
XX 20-MAR-2003.
XX PF
XX 09-SEP-2002; 2002WO-US028539.
XX PR
XX 07-SEP-2001; 2001US-0318120P.
XX PR
XX 07-SEP-2001; 2001US-0318130P.
XX PR
XX 10-SEP-2001; 2001US-0318430P.
XX PR
XX 17-SEP-2001; 2001US-0322636P.
XX PR
XX 17-SEP-2001; 2001US-0322781P.
XX PR
XX 17-SEP-2001; 2001US-0322816P.
XX PR
XX 17-SEP-2001; 2001US-0322817P.
XX PR
XX 19-SEP-2001; 2001US-0323519P.
XX PR
XX 20-SEP-2001; 2001US-0323631P.
XX PR
XX 20-SEP-2001; 2001US-0323636P.
XX PR
XX 25-SEP-2001; 2001US-0324969P.
XX PR
XX 26-SEP-2001; 2001US-0325091P.
XX PR
XX 26-SEP-2001; 2001US-0324990P.
XX PR
XX 17-APR-2002; 2002US-0373212P.
XX PR
XX 06-SEP-2002; 2002US-00236177.
XX (CURA-) CURAGEN CORP.
XX PA
XX Spytek KA, Patturajan M, Gorman L, Li L, Anderson DW, Zhong M;
XX Gerlach VL, Vernet CAM, Ellerman K, Berghs C, Rothenberg ME, Guo X;
XX Shinkets RA, Leach MD, Catterton E, Kekuda R, Ji W, Miller CE;
XX Rieger DK, Taupier RJ, Shenoy SG, Liu X, Padigar M, Aleobrook JP;
XX Lepley DM, Edinger SR, Borgees CE;
XX WPI; 2003-313242/30.
XX P-PSDB; ABO14747.
XX
XX New cytoplasmic, nuclear membrane bound or secreted polypeptides (NOVX)
XX and polynucleotides, useful in gene therapy, e.g. for treating or
XX preventing obesity, multiple sclerosis, allergy, cancers, hemophilia,
XX stroke or infections.
XX Claim 20; Page 336-340; 586pp; English.
XX The invention describes a new isolated polypeptide (NOVX). The NOVX
XX polypeptide, nucleic acid and antibody are useful as therapeutics,
XX particularly in the manufacture of a medicament for treating a syndrome

CC associated with a human disease, which includes a pathology associated
CC with NOVX polypeptide. The DNA encoding the protein is useful in gene
CC therapy for treating the disease or condition. In particular, the NOVX
CC polypeptide or polynucleotide is useful for treating endocrine/
CC metabolism-related diseases (e.g. obesity or diabetes), central nervous
CC system disorders (e.g. Alzheimer's disease, Parkinson's disease,
CC epilepsy, multiple sclerosis, schizophrenia or depression), autoimmune
CC and inflammatory disorders (e.g. psoriasis, allergy, lupus erythematosus,
CC asthma, inflammatory bowel disease, rheumatoid arthritis or
CC osteoarthritis), cancers (e.g. colon, lung, liver, breast, ovarian,
CC prostate or brain cancers, or melanoma), liver diseases (e.g. liver
CC cirrhosis), lung diseases (emphysema or obstructive pulmonary disease),
CC haemophilia, stroke, or infections (e.g. viral, bacterial or parasitic).
CC These are also useful in developing powerful assay system for functional
CC analysis of various human disorders, as well as in diagnostic
CC applications, and for monitoring the effects of drugs during clinical
CC trials. This sequence encodes a novel human NOV protein
XX
SQ Sequence 14994 BP; 4902 A; 3673 C; 3081 G; 3338 T; 0 U; 0 Other;

Query Match 10.1%; Score 41; DB 7; Length 14994;
Best Local Similarity 48.5%; Pred. No. 0.33;
Matches 113; Conservative 0; Mismatches 120; Indels 0; Gaps 0;
Qy 62 AAGGAAACCCAGATGAGGAAAGAAACCGCAGGAAATGACTACATTTGAAAGAAACTTC 121
Db 4160 ATGAAAGTCAGAAAGAAACACACACCCCATGAAATTTCTCTGACACGCTAAGACC 4219
Qy 122 AAGATCAAGATAAGAAAGCCAGAAAGTTTCATCCACTTCTATATCAGGAAACGAGATG 181
Db 4220 AAGAGAAACCTCAGAGTTTATCTGAAACCTTGAAATTTACTATTTCAGAGGAGATCA 4279
Qy 182 GCAGTGGTTCTGAGAAAGTGCTACACTGTCTATTATCATCCCCCATCAGAGATCCT 241
Db 4280 AAGAGATCAAGAAAGAAAGGAAAGACACTTTTAAAGTAGCCACAGATATTCCTT 4339
Qy 242 CCCTGAGCTCCAAATGATGANGGCTATGAGAACATTGACTCCCTCACAAGGAAA 294
Db 4340 CCAGCAGGACCAATAAGAGAAAGTCTGAGTTTGTGATGACATAACTACTAGA 4392

Search completed: April 22, 2004, 14:55:21
Job time : 359 secs

[illegible]

121	Qy	CAAGATCAAGATAGAGAAAGCCAGAGAGTTTCATCTTCTAATCAGGAAACAGAGAA	180
145	Db	CAAGATCAAGATAGAGAAAGCCAGAGAGTTTCATCTTCTAATCAGGAAACAGAGAA	204
181	Qy	GGCAGTGGTTCGAGAAAGTGTGTACACTGTCTAATTAATCACATATCCCATCAGAGATCC	240
205	Db	GGCAGTGGTTCGAGAAAGTGTGTACACTGTCTAATTAATCACATATCCCATCAGAGATCC	264
241	Qy	TCCTTGAGTCCAAATGATGATGGCTATAGAGAACTTGACTCCCTCTCAAGGAAAGTGAGA	300
265	Db	TCCTTGAGTCCAAATGATGATGGCTATAGAGAACTTGACTCCCTCTCAAGGAAAGTGAGA	324
301	Qy	CAGTTTTAGAGAAAGGTACAGACAGCAATATGCCCTCTTTAGGACTTCTGTGTAGTAGCCCT	360
325	Db	CAGTTTTAGAGAAAGGTACAGACAGCAATATGCCCTCTTTAGGACTTCTGTGTAGTAGCCCT	384
361	Qy	TGTTCTGCACCCATGAGCATGATTATGAAGTTGTGTTTCCACAC	405
385	Db	TGTTCTGCACCCATGAGCATGATTATGAAGTTGTGTTTCCACAC	429

RESULT 2

US-10-221-714A-50/c
; Sequence 50, Application US/10221714A
; Publication No. US20040048254A1

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

Query Match

Query Match 10.3%; Score 41.6; DB 13; Length 6446;
Best Local Similarity 56.6%; Pred.No. 0.31;
Matches 77; Conservative 0; Mismatches 59; Indels 0;

RESULT 4

US-10-311-455-673/c
; Sequence 673, Application US/10311455
; Publication No. US20030143606A1

RESULT 3

US-10-221-714A-63/c
; Sequence 63, Application US/10221714A
; Publication No. US20040048254A1

[illegible]

RESULT 4

RESOL 4
US-10-311-455-673/c
; Sequence 673, Application US/10311455
; Publication No. US20030143606A1

	Matches	125;	Conservative	0;	Mismatches	139;	Indels	0;	Gaps	0;
QY	39	AGAGAATCAAAAGAACGCCAAGAACCCGACATGAGGAAGAAACAACCGCAGGAAT	98							
Db	18346	AAAAATTAAAATAATTATATCAAATAATAACTTTATCAACAACATAAACCCCAAAAAA	18287							
QY	99	GACTACATTTGAAAGAAAACCTTCAAGATCAAGATAAGAAAAGCCCAAGAAAGTTTCATCCAC	158							
Db	18286	AACAATTTACAAAAAANAACCAAAATTCACACACCTTAATCAACAACAATTTTCCTC	18227							
QY	159	TCTAATCAGGAAAAACGAGATGGCAGTGTTCTGAAGAAGTGCTCACATGTCAATTA	218							
Db	18226	TCTAAATCAATAAAAAACAACCTTACCACCTATATCAATATTTATCACCCTACTATCA	18167							
QY	219	TCACATCCCCCATCAGAGATCCTCCCTCAGCTCCAATCATGATGCCTATGAGACATTGA	278							
Db	18166	CAACCAATTATTCACAAATATCTCCTATCTTCAAAACATCGAACACCTATCACCTTAA	18107							
QY	279	CTCCCTCAACAAGAAAGTCAGACA	302							
Db	18106	ACCCCTCCATAAAAAATAACACA	18083							

RESULT 6
US-09-864-761-27195/c
; Sequence 27195, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USED FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aomics-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/532,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24253.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/508,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29

Db 735 TGGATGATTATATAGAGAACTCACTCAGTTCTTCTGCTAAAAAAGAAAAAAGAAAAA 794
QY 62 AAGGAACCCAGATGAGGAAGAAACCGCAGGAAATGACTACATTGGAAGAACTTC 121
Db 795 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 854
QY 122 AAGATCAAGATAGAAAGAAAGCCAAAGAA 147
Db 855 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 880

RESULT 9

US-09-925-301-223
; Sequence 223, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 223
; LENGTH: 2921
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1609)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (2919)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (2920)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (2921)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-301-223

Query Match 9.9%; Score 40; DB 9; Length 2921;
Best Local Similarity 54.1%; Pred. No. 0.63;
Matches 79; Conservative 1; Mismatches 66; Indels 0; Gaps 0;
QY 2 TGGGAATTTATCTCTCGGAAACTCAGTTGCTGGGAGAGAAATCAAGAGCCCAAGA 61
Db 2742 TGGGAATTTAAATCCATTAATACATTTTGAAGTGTGWAAGAAAAAAGAAAAA 2801
QY 62 AAGGAACCCAGATGAGGAAGAAACCGCAGGAAATGACTACATTGGAAGAACTTC 121
Db 2802 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 2861
QY 122 AAGATCAAGATAGAAAGCCAAAGAA 147
Db 2862 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 2887

RESULT 10

US-09-925-301-307
; Sequence 307, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10

; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 307
; LENGTH: 785
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-301-307

Query Match 9.8%; Score 39.8; DB 9; Length 785;
Best Local Similarity 54.4%; Pred. No. 0.39;
Matches 80; Conservative 0; Mismatches 67; Indels 0; Gaps 0;
QY 1 ATGGGAATTTATCTCTCGGAAACTCAGTTGCTGGGAGAGAAATCAAGAGCCCAAG 60
Db 629 ATTGGTATACCTTATACAGAAAGCCCAATAGACCCAAAGTCTCCAAAGAACTCAAG 688
QY 61 AAGGAACCCAGATGAGGAAGAAACCGCAGGAAATGACTACATTGGAAGAACTTC 120
Db 689 GAAGAGAAACGAAACACAGAGCAAGAAATATAATAATAATAATAATAATAATAA 748
QY 121 CAAGATCAAGATAGAAAGCCCAAGAA 147
Db 749 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 775

RESULT 11

US-10-087-192-913/c
; Sequence 913, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 52945200122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 913
; LENGTH: 56258
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)....(56258)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-913

Query Match 9.8%; Score 39.8; DB 13; Length 56258;
Best Local Similarity 60.7%; Pred. No. 2.8;
Matches 65; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
QY 40 GGAATCAAGAGAGCCCAAGAAAGAAACCCAGATGAGGAAGAAACCGCAGGAATG 99
Db 11067 GAGAAGGAGAGGAGAGGAGAGGAGAAAGGAAAGGAAAGGAAAGGAAAGGAAAG 11008
QY 100 ACTACATTTTGAAGAAAGAACTTCAGATCAAGATAGAAAGCCCAAGAA 146
Db 11007 AAGGAAAGGAGAGGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 10961

RESULT 12

US-10-335-977-4490
; Sequence 4490, Application US/10335977

```
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
;
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD/ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
; SOFTWARE: UNIX
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/335,977
; FILING DATE: 30-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/993,002
; FILING DATE: 17-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 742-4214
; INFORMATION FOR SEQ ID NO: 4490:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 489 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...489
; SEQUENCE DESCRIPTION: SEQ ID NO: 4490:
US-10-335-977-4490
Query Match 9.7%; Score 39.2; DB 13; Length 489;
Best Local Similarity 64.1%; Pred. No. 0.47;
Matches 59; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 57 CAAGAAAGGAAACCCAGATGAGGAAGAAACGGCAGGAATGACTACATTGGAAGAAA 116
Db 108 CAATGAACATACTCTACAGAGAAAAGTTAGAGCCACGCTTAATTAGAAGAAA 167

QY 117 ACTTCAAGATCAAGATAAGAAAAGCCAAAG 148
Db 168 ACACAAAGAACCAAGCAAGCAAGCAAG 199

RESULT 13
US-10-335-977-4492
; Sequence 4492, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
;
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
```

```
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD/ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
; SOFTWARE: UNIX
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/335,977
; FILING DATE: 30-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/993,002
; FILING DATE: 17-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 742-4214
; INFORMATION FOR SEQ ID NO: 4492:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 963 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...963
; SEQUENCE DESCRIPTION: SEQ ID NO: 4492:
US-10-335-977-4492
Query Match 9.7%; Score 39.2; DB 13; Length 963;
Best Local Similarity 64.1%; Pred. No. 0.65;
Matches 59; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 57 CAAGAAAGGAAACCCAGATGAGGAAGAAACGGCAGGAATGACTACATTGGAAGAAA 116
Db 582 CAATGAACATACTCTACAGAGAAAAGTTAGAGCCACGCTTAATTAGAAGAAA 641

QY 117 ACTTCAAGATCAAGATAAGAAAAGCCAAAG 148
Db 642 ACACAAAGAACCAAGCAAGCAAGCAAG 673

RESULT 14
US-10-335-977-4491
; Sequence 4491, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
;
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD/ROM ISO9660
```

COMPUTER: IBM PC Compatible
OPERATING SYSTEM: Windows NT 4.0
SOFTWARE: UNIX
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/335,977
FILING DATE: 30-Dec-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/993,002
FILING DATE: 17-DEC-1997
ATTORNEY/AGENT INFORMATION:
NAME: Mandragoras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: GYN-018
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 4491:
SEQUENCE CHARACTERISTICS:
LENGTH: 966 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Helicobacter pylori
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...966
SEQUENCE DESCRIPTION: SEQ ID NO: 4491:
US-10-335-977-4491

Query Match 9.7%; Score 39.2; DB 13; Length 966;
Best Local Similarity 64.1%; Pred. No. 0.65;
Matches 59; Conservative 0; Mismatches 33; Indels 0; Gaps 0;
QY 57 CAAGAAAGGAACCCAGATGAGGAAAGAAACGGCAGGAAATGACTACATTTCGAAGAA 116
DB 582 CAATGAACACTAATCTCAGAGAGAAAAAAGTTAGAGCCACGGCTTAATTAGAGAAA 641
QY 117 ACTTCAAGATCAAGATAGAAAGCCCAAGAG 148
DB 642 ACACAAAGAACACAGACAGCAAGCAAGCAAGAG 673

RESULT 15
US-10-311-455-78/c
Sequence 78, Application US/10311455
Publication No. US20030143606A1
GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: PIEPENBROCK, Christian
APPLICANT: BERLIN, Kurt
TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining Cytosine Methylation
FILE REFERENCE: 5013.1014
CURRENT APPLICATION NUMBER: US/10/311.455
CURRENT FILING DATE: 2002-12-16
PRIOR APPLICATION NUMBER: PCT/EP01/07537
PRIOR FILING DATE: 2001-07-02
PRIOR APPLICATION NUMBER: DE 10032529.7
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: DE 10043826.1
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 2424
SEQ ID NO 78
LENGTH: 17869
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
FEATURE:

NAME/KEY: unsure
LOCATION: 3465, 3586
OTHER INFORMATION: n is a or g or c or t
US-10-311-455-78
Query Match 9.7%; Score 39.2; DB 15; Length 17869;
Best Local Similarity 47.9%; Pred. No. 2.5;
Matches 113; Conservative 0; Mismatches 123; Indels 0; Gaps 0;
QY 6 AAATTATCTCTCGGAAACTCAGTTGCTGGGAGAGATCAAAAGAGAGCCCAAGAAAGG 65
DB 8214 AAACATATCCAAAACTAATCTTAATAATAATAATAATAATAATAATAATAATAA 8155
QY 66 AAACCCAGATGAGGAAAGAAACCGCAGGAAATGACTACATTTCGAAGAAAGAACTTCAGA 125
DB 8154 AAACCTAAATACAAACTTAAACCAAAATCGACTAACGAATTTTAAAAAAATCACTAAA 8095
QY 126 TCAGATTAAGAAAGCCCAAGAGTTTCATCCACTTCTAATCAGGAAACGAGAGATGGCAG 185
DB 8094 AATCTTAAACAAAAAAAATTAATTTCAACGGCAAACTAATTATATATATATATAC 8035
QY 186 TGGTTCTGAAGAGTGTGCTACACTGTGCTAATTAATCACAATCCCCATCAGAGATCCT 241
DB 8034 TTTCACTTAAAAAATATTTTAACTTCTTCCAACTAAACAAATATTTTATTACT 7979
Search completed: April 22, 2004, 17:14:09
Job time : 347 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 22, 2004, 13:49:27 ; Search time 80 seconds
(without alignments)

2809.439 Million cell updates/sec

Title: US-10-005-907-1_COPY_25_429

Perfect score: 405

Sequence: 1 atgggaattatctctgog.....atgaagtgtgtttccacac 405

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents NA:*
- 1: /cgm2_6/ptodata/2/ina/5A COMB.seq.*
 - 2: /cgm2_6/ptodata/2/ina/5B COMB.seq.*
 - 3: /cgm2_6/ptodata/2/ina/6A COMB.seq.*
 - 4: /cgm2_6/ptodata/2/ina/6B COMB.seq.*
 - 5: /cgm2_6/ptodata/2/ina/PCUTS COMB.seq.*
 - 6: /cgm2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	49	12.1	7218	1	US-08-232-463-14
C 2	38.2	9.4	1092	4	US-08-328-352-1567
C 3	36.8	9.1	549	4	US-09-107-532A-2987
C 4	36.4	9.0	580073	4	US-08-545-528D-1
C 5	35.4	8.7	2608	4	US-09-904-615-16
C 6	35	8.6	759	4	US-09-540-236-1733
C 7	35	8.6	2447	2	US-09-014-969-14
C 8	34.8	8.6	399	4	US-09-621-976-8976
C 9	34.8	8.6	729	4	US-09-134-001C-1161
C 10	34.8	8.6	118067	4	US-09-497-855A-32
C 11	34	8.4	1462	4	US-09-133-862A-5
C 12	34	8.4	1985	4	US-09-907-794A-212
C 13	34	8.4	1985	4	US-09-905-125A-212
C 14	34	8.4	1985	4	US-09-902-775A-212
C 15	33.8	8.3	5340	4	US-09-627-122-21
C 16	33.6	8.3	249	4	US-09-621-976-1322
C 17	33.6	8.3	3275	4	US-09-370-838-151
C 18	33.4	8.2	240	4	US-09-621-976-1324
C 19	33.4	8.2	640	4	US-09-976-594-1142
C 20	33.2	8.2	926	3	US-08-945-994-8
C 21	33.2	8.2	5727	4	US-08-956-171E-190
C 22	33	8.1	400	4	US-08-956-171E-4053
C 23	33	8.1	674	4	US-09-620-405B-465
C 24	33	8.1	674	4	US-09-433-826B-465
C 25	33	8.1	674	4	US-09-604-287A-465
C 26	33	8.1	674	4	US-09-834-759-465
C 27	33	8.1	1000	3	US-09-018-584A-34

Sequence 108, Appl
Sequence 3, Appl
Sequence 33, Appl
Sequence 9, Appl
Sequence 9, Appl
Sequence 9, Appl
Sequence 9, Appl
Sequence 9, Appl
Sequence 9, Appl
Sequence 9, Appl
Sequence 9, Appl
Sequence 15, Appl
Sequence 41, Appl
Sequence 45, Appl
Sequence 41, Appl
Sequence 45, Appl
Sequence 38, Appl
Sequence 38, Appl
Sequence 17, Appl
Sequence 13, Appl

ALIGNMENTS

RESULT 1
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZ9pt-Fis
; US-08-232-463-14

Query Match 12.1%; Score 49; DB 1; Length 7218;

Best Local Similarity 2.2%; Pred. No. 4e-05;
Matches 7; Conservative 192; Mismatches 122; Indels 0; Gaps 0;
QY 7 AATTATCTCTCGGAAACTCAGTTGCTGGAGAGATCAAAAGAGAGCCCAAGAGGA 66
Db 1447 AAGAAATTTGGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1388
QY 67 AACCCAGATGAGGAAGAAACGGCAGGAATGACTACTATTGAAAGAAACTTCAAGAT 126
Db 1307 RRR 1328
QY 127 CAAGTAAGAAAGCAAGAGTTTCATCCACTTCTAATCAGGAAACGAGATGGCAGT 186
Db 1327 RRR 1268
QY 187 GCTTCTGAGAAAGTGTGACATGTCATTAATCATCATCCCTCCATCAGATCTCCCTG 246
Db 1267 RRR 1208
QY 247 AGCTCAATGATGCTATGAGAACATGACTCCCTCAAGAAAGTGAAGACATTT 306
Db 1207 RRR 1148
QY 307 AGAGAAAGTCAAGACAGAA 327
Db 1147 RRRRRRRRRRRRRRRRRRR 1127

RESULT 2
US-09-328-352-1567/c
; Sequence 1567, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTG99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 1567
; TYPE: DNA
; LENGTH: 1092
; ORGANISM: Acinetobacter baumannii
US-09-328-352-1567

Query Match 9.4%; Score 38.2; DB 4; Length 1092;
Best Local Similarity 52.9%; Pred. No. 0.037; 73; Indels 0; Gaps 0;
Matches 82; Conservative 0; Mismatches 73; Indels 0; Gaps 0;
QY 43 AATCAAAAGAGCCCAAGAAAGGAAACCCAGATGAGGAAAGAAACCGCAGGAATGACT 102
Db 315 AATAATTAGAAACCAATAAACAGGCTTCAGAGGTGATACCCACGGGTGTAATGGGT 256
QY 103 ACATTTGAAGAAACTTCAGATCAAGATAGAAAGCCAGAGTTTCATCCACTTCT 162
Db 255 ACGGTAGAAAGACAATCTTGGGGCAATGATATGTTTCAAGAACCATGCCCCAATACT 196
QY 163 AATCAGAAACGAGATGGCAGTGGTTCTGAAGA 197
Db 195 GCTCATTAAGAAAGAAATTTGGCTGCAAGACAGACA 161

RESULT 3
US-09-107-532A-2987
; Sequence 2987, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:

ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneka
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 2987:
SEQUENCE CHARACTERISTICS:
LENGTH: 549 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...549
SEQUENCE DESCRIPTION: SEQ ID NO: 2987:
US-09-107-532A-2987

Query Match 9.1%; Score 36.8; DB 4; Length 549;
Best Local Similarity 50.0%; Pred. No. 0.072;
Matches 92; Conservative 0; Mismatches 92; Indels 0; Gaps 0;
QY 35 TGGGAGAGATCAAAAGAGCCCAAGAGGAAACCCAGATGAGGAAAGAAACCGCAGG 94
Db 29 TTGGGAACAATAAGAGACTCGCCAAATAAATAAAGAGAGAGCAAGCAAGAAAGAGC 88
QY 95 AATGACTACATTTGAAAGAAACTTCAAGATCAAGATAAGAAAGCCCAAGAGTTTCTAT 154
Db 89 AAGCAATTACCATCTGAAATAAGCAATTAGAAAGTGCCACCAAGTTCAAGAAATAAAA 148
QY 155 CCATTTCTAATCAGGAAACGAGATGGCAGTGGTTCTGAGAAAGTGTGCTACTGTCA 214
Db 149 CTTCTCCCAATTTGAAAGACAGAGTAGAGGAAGAACAAAGAAAGTTGGCGCCCAATCA 208
QY 215 TTAA 218
Db 209 GAA 212

RESULT 4
US-08-545-528D-1/c
; Sequence 1, Application US/08545528D
; Patent No. 6537773
; GENERAL INFORMATION:
; APPLICANT: Fraser et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Mycoplasma Genitium Genome, Fragments
; Patent No. 6537773
; TITLE OF INVENTION: Thereof, and Uses Thereof

FILE REFERENCE: PB193P1
CURRENT APPLICATION NUMBER: US/08/545,528D
CURRENT FILING DATE: 1995-10-19
PRIOR APPLICATION NUMBER: US 08/488,018
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: US 08/473,545
PRIOR FILING DATE: 1995-06-07
NUMBER OF SEQ ID NOS: 1
SOFTWARE: Patent in version 3.1
SEQ ID NO 1
TYPE: DNA
ORGANISM: Mycoplasma genitalium.
US-08-545-528D-1

Query Match 9.0%; Score 36.4; DB 4; Length 580073;
Best Local Similarity 51.9%; Pred. No. 2.4;
Matches 82; Conservative 0; Mismatches 76; Indels 0; Gaps 0;
QY 37 GGAGAGAATCAAAGAGCCCAAGAGGAAACCCAGATGAGGAAAGAAACGGCAGGAA 96
Db 502562 GAAGTTGATATTATGAACTTAAAGAAAGAAACAGAACTTGAAGAAATTCACAT 502503
QY 97 ATGACTACATTTGAAGAAACTTCAAGATCAAGATAAGAAAGCCCAAGATTCATCC 156
Db 502502 ATCAAGGATGCTAAGCTAAACGTTAAATTTGAACAAATTTAATGTTGAACCTATTAAAA 502443
QY 157 ACTTCTAATCAGGAACGAGATGCGAGTGGTCTGA 194
Db 502442 CTTTCAGTAAGAAATAAATAATGCTGTTTCTTAA 502405

RESULT 5
US-09-904-615-16
Sequence 16, Application US/09904615
Patent No. 6566325
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 49 Human Secreted Proteins
FILE REFERENCE: P2032P1
CURRENT APPLICATION NUMBER: US/09/904,615
CURRENT FILING DATE: 2001-07-16
PRIOR APPLICATION NUMBER: 09/511,554
PRIOR FILING DATE: 2000-02-23
PRIOR APPLICATION NUMBER: 60/097,917
PRIOR FILING DATE: 1998-08-25
PRIOR APPLICATION NUMBER: 60/098,634
PRIOR FILING DATE: 1998-08-31
NUMBER OF SEQ ID NOS: 170
SOFTWARE: Patent in Ver. 2.0
SEQ ID NO 16
LENGTH: 2608
TYPE: DNA
ORGANISM: Homo sapiens
US-09-904-615-16

Query Match 8.7%; Score 35.4; DB 4; Length 2608;
Best Local Similarity 54.1%; Pred. No. 0.4;
Matches 72; Conservative 0; Mismatches 61; Indels 0; Gaps 0;
QY 15 CCTGCGAAACTCAGTTCCTGGGAGAGATCAAAAGAGCCCAAGAGAAAGAAACCCAGA 74
Db 2474 CTTTCAATAACAGCTCCTGGGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 2533
QY 75 TGAGGAAGAAACGGCAGAGATGATCATTTGAAGAAAGAACTTCAGATCAAGATTA 134
Db 2534 AA 2593
QY 135 GAAAGCCAGAA 147
Db 2594 AAAAAAAAAAAAA 2606

RESULT 6
US-09-540-236-1733
Sequence 1733, Application US/09540236
Patent No. 6673910
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR
FILE REFERENCE: 2709.2005-001
CURRENT APPLICATION NUMBER: US/09/540,236
CURRENT FILING DATE: 2000-04-04
NUMBER OF SEQ ID NOS: 3840
SEQ ID NO 1733
LENGTH: 759
TYPE: DNA
ORGANISM: M. catarrhalis
US-09-540-236-1733

Query Match 8.6%; Score 35; DB 4; Length 759;
Best Local Similarity 48.3%; Pred. No. 0.3;
Matches 98; Conservative 0; Mismatches 105; Indels 0; Gaps 0;
QY 159 TTCTAATCAGGAAACGAGATGCGAGTGTCTGAAGAGTGTGTACATGTCATTAA 218
Db 171 TTATGATAATAATCGTTGGTATATCGAAGGTCTGAAGCGGGTTTTTACCCCTTAAGGA 230
QY 219 TCACATCCCCCATCAGAGATCCTCCCTGAGCTCCAATGATGATGCTATGAGAACATTGA 278
Db 231 TAACAGACCATGTCGTATCGCTTAACATTATGATGGTCAGAGCTTGGACCCCATCA 290
QY 279 CTCCCTCAAGGAAAGTGAGACAGATTAGAGAAGGTGAGAGACAGAAATATGCCCTTCT 338
Db 291 TGCCAATGCCATAGAACTTAAGCAGCTTAATAAGCGTCAATGCTCAGCATCGGCCCATGC 350
QY 339 TAGGACTTCTGTAGTAGGCCTT 361
Db 351 AGCTATCTTTATATCACCCTT 373

RESULT 7
US-09-014-969-14
Sequence 14, Application US/09014969
Patent No. 5965397
GENERAL INFORMATION:
APPLICANT: Jacobs, Kenneth
APPLICANT: McCoy, John M.
APPLICANT: LaValle, Edward R.
APPLICANT: Racie, Lisa A.
APPLICANT: Merberg, David
APPLICANT: Treacy, Maurice
APPLICANT: Spaulding, Vikki
APPLICANT: Agostino, Michael J.
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridge Park Drive
CITY: Cambridge
STATE: MA
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/014,969
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Sprunger, Suzanne A.

REGISTRATION NUMBER: 41,323
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8284
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 2447 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-014-969-14

Query Match 8.6%; Score 35; DB 2; Length 2447;
Best Local Similarity 50.9%; Pred. No. 0.52;
Matches 83; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 17 TCGGAAACTCAGTTGCTGGGAGAGATCAAAAGAGCCCAAGAGGAAACCCAGATG 76
DB 2221 TGTGCAATAAACCTTTTGGGAGAAAAAARARARARARARARARARARARAR 2280

QY 77 AGGAAAGAAACCGCAGGAATGACTACATTTGAAAGAAACCTTCAAGATCAAGATGA 136
DB 2281 AA 2340

QY 137 AAGCCAGAGATTTCATCCACTTCTAATCAGGAAACGAGAA 179
DB 2341 AA 2383

RESULT 8
US-09-621-976-8976
Sequence 8976, Application US/09621976
Patent No. 6639063
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Jobert, S.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: GENSET.054PR2
CURRENT APPLICATION NUMBER: US/09/621,976
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.pm
SEQ ID NO 8976
LENGTH: 399
TYPE: DNA
ORGANISM: Homo sapiens
US-09-621-976-8976

Query Match 8.6%; Score 34.8; DB 4; Length 399;
Best Local Similarity 12.4%; Pred. No. 0.26;
Matches 30; Conservative 110; Mismatches 102; Indels 0; Gaps 0;

QY 3 GGGAAATTATCTCTCGGAAACTCAGTTGCTGGGAGAGATCAAAAGAGCCCAAGAA 62
DB 42 KRRRRRRANWKKSCWCKSKSWRSWGWTKEMKGRGAASWAGYMSWNTYTRWRY 101

QY 63 AGGAAACCGATGAGGAAAGAAACCGCAGGAAATGACTACATTTGAAAGAAACCTTCA 122
DB 102 YRYRKACTKWRAGCWGKAGWAWAYAKWYMWYMWRTAMKYWAMKSKRSRRRRAWYA 161

QY 123 AGATCAAGATGAAGAAACCGCAAGAGTTTCATCCACTTCTAATCAGGAAACGAGATGG 182
DB 162 WNYWYARTTWGTPASCYRGHYMASAGYNYWYMWYMRKWKWYSGWSMKWTRCAS 221

QY 183 CAGTGGTTCTGAGAGTGTGTACACTGTCTAATCAATCAATCCATCCATCAGATCCTC 242
DB 222 YSCWSSYCWGAKWMMYKTSRWSYWSYRCKYRRCSCCWSSNSCYWKTYYRWSYCAS 281

QY 243 CC 244
DB 282 Y 283

RESULT 9
US-09-134-001C-1161
Sequence 1161, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 1161
LENGTH: 729
TYPE: DNA
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-1161

Query Match 8.6%; Score 34.8; DB 4; Length 729;
Best Local Similarity 49.5%; Pred. No. 0.34;
Matches 90; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

QY 40 GAGATCAAAAGAGCCCAAGGAAAGCAAGCAAGATGAGGAAAGCAAGCAAGCAAGCAAGATG 99
DB 358 GAACACGTAAACAAAGAAAGAAACCATCAAGATAAACAAAGACGTCGCGATAAG 417

QY 100 ACTACATTTGAAAGAAACTTCAAGATCAAGATCAAGATCAAGATCAAGATCAAGATCAAGAT 159
DB 418 CCATCAATGTGAGAAATATTATAGTGAAGTTTGACGATTTAGAGCAAGTTATCGACTAT 477

QY 160 TCTATTCAGGAAACGAGATGCGCAGTGTCTTGAAGAGTGTGCTACACTGTCTATTAAT 219
DB 478 GCTTATCATTAATCAAAACACTGATGAATTTGAAGATTTATTATATATATGATGATAAT 537

QY 220 CA 221
DB 538 AA 539

RESULT 10
US-09-497-855A-32/c
Sequence 32, Application US/09497855A
Patent No. 6605432
GENERAL INFORMATION:
APPLICANT: Huang, Tim
TITLE OF INVENTION: HIGH-THROUGHPUT METHODS FOR DETECTING DNA METHYLATION
FILE REFERENCE: UMO1523
CURRENT APPLICATION NUMBER: US/09/497,855A
CURRENT FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: 60/120,592
PRIOR FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: 60/118,760
PRIOR FILING DATE: 1999-02-05
NUMBER OF SEQ ID NOS: 54
SOFTWARE: PatentIn version 3.0
SEQ ID NO 32
LENGTH: 118067
TYPE: DNA
ORGANISM: Homo sapiens;
US-09-497-855A-32

Query Match 8.6%; Score 34.8; DB 4; Length 118067;
Best Local Similarity 53.7%; Pred. No. 3.5;
Matches 72; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 37 GGAGAGATCAAAAGAGCCCAAGGAAAGCAAGATGAGGAAAGCAAGCAAGCAAGCAAGCAAG 96


```
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 212
; LENGTH: 1985
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-907-794A-212

Query Match      8.4%; Score 34; DB 4; Length 1985;
Best Local Similarity 52.1%; Pred. No. 0.95;
Matches 76; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 2 TGGGAATATCTCTCGGAAACTCAGTTGCTGGGAGAGATCAAAAGAGCCCAAGA 61
Db 1824 TTGACATATAATGGTGTATGCTTCCGCCAAATAAAAAAAAAAAAAAAAAA 1883
QY 62 AAGGAAACCCAGATGAGGAAAGAAACGGCAGGAAATGACTACATTTCAAGAAACTTC 121
Db 1894 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1943
QY 122 AAGATCAAGATAGAAAGCCCAAGA 147
Db 1944 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1969
```

RESULT 13

```
US-09-905-125A-212
; Sequence 212, Application US/09905125A
; Patent No. 6664376
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gieritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kijavrin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/905,125A
; CURRENT FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
```

```
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 212
; LENGTH: 1985
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-905-125A-212

Query Match      8.4%; Score 34; DB 4; Length 1985;
Best Local Similarity 52.1%; Pred. No. 0.95;
Matches 76; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 2 TGGGAATATCTCTCGGAAACTCAGTTGCTGGGAGAGATCAAAAGAGCCCAAGA 61
Db 1824 TTGACATATAATGGTGTATGCTTCCGCCAAATAAAAAAAAAAAAAAAAAA 1883
QY 62 AAGGAAACCCAGATGAGGAAAGAAACGGCAGGAAATGACTACATTTCAAGAAACTTC 121
Db 1884 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1943
QY 122 AAGATCAAGATAGAAAGCCCAAGA 147
Db 1944 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1969
```

RESULT 14

```
US-09-902-775A-212
; Sequence 212, Application US/09902775A
; Patent No. 6686451
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gieritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
```

APPLICANT: Kljavin, Ivar J.
APPLICANT: Mather, Jennie P.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William, I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
TITLE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: 10466-14
CURRENT APPLICATION NUMBER: US/09/902,775A
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2001-07-10
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: US/60/143,048
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: US/60/145,698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: US/60/146,222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/20944
PRIOR FILING DATE: 1999-09-13
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21547
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/23089
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: PCT/US99/28214
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: PCT/US99/28564
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/28565
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: PCT/US99/30911
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US99/30999
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US00/00219
PRIOR FILING DATE: 2000-01-05
NUMBER OF SEQ ID NOS: 423
SEQ ID NO 212
LENGTH: 1985
TYPE: DNA
ORGANISM: Homo sapiens
US-09-902-775A-212

Query Match 8.4%; Score 34; DB 4; Length 1985;
Best Local Similarity 52.1%; Pred. No. 0.95; Indels 0; Gaps 0;
Matches 76; Conservative 0; Mismatches 70; Indels 0; Gaps 0;
QY 2 TGGGAATTTATCTCTCGGAAATCTCAGTTGCTGGGAGAGATCAAAAGAGCCCAAGA 61
DB 1824 TTGGACAATAAATGGTGCTATGACTGCTTCGGCCAAAAAATAAAAAAAAAAAAA 1883
QY 62 AAGGAACCCAGATCAG 121
DB 1884 AA 1943
QY 122 AAGATCAAGATAAGAAAGCCCAAGAA 147
DB 1944 AA 1969

RESULT 15

US-09-627-122-21/c
Sequence 21, Application US/09627122
Patent No. 6472521
GENERAL INFORMATION:
APPLICANT: Uhlmann, Eugen
APPLICANT: Greiner, Beate
APPLICANT: Unger, Eberhard
APPLICANT: Gothe, Gislinde
APPLICANT: Schwerdel, Marc
TITLE OF INVENTION: OLIGONUCLEOTIDES FOR THE INHIBITION OF HUMAN eg5
FILE REFERENCE: 02481.1678
CURRENT APPLICATION NUMBER: US/09/627,122
CURRENT FILING DATE: 2000-07-27
NUMBER OF SEQ ID NOS: 23
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 21
LENGTH: 5340
TYPE: DNA
ORGANISM: Plasmodium falciparum
US-09-627-122-21
Query Match 8.3%; Score 33.8; DB 4; Length 5340;
Best Local Similarity 49.7%; Pred. No. 1.7;
Matches 86; Conservative 0; Mismatches 87; Indels 0; Gaps 0;
QY 1 ATGGGAATTTATCTCTCGGAAATCTCAGTTGCTGGGAGAGATCAAAAGAGCCCAAG 60
DB 4840 ATGGGAATTTCTTTGCTCCCATGTTAATTTTACACATATATATATATATATATAT 4781
QY 61 AAGGAACCCAGATCAG 120
DB 4780 AAAAAATAATATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAAT 4721
QY 121 CAAGATCAAGATAAGAAAGCCCAAGAGTTTCATCCACTTCTAATCAGGAAAA 173
DB 4720 AAAAAATAATACATATACATATATATATATATATATATATATATATATATATAT 4668
Search completed: April 22, 2004, 16:09:14
Job time : 86 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 22, 2004, 18:31:30 ; Search time 2076 Seconds
(without alignments)
8455.650 Million cell updates/sec

Title: US-10-005-907-1_COPY_25_429
Perfect score: 405
Sequence: 1 atgggaattatctctgcg.....atgaagtgtgtttccacac 405

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 3364370

Minimum DB seq_length: 0
Maximum DB seq_length: 405

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:**

- 1: gb_ba:**
- 2: gb_htg:**
- 3: gb_in:**
- 4: gb_om:**
- 5: gb_ov:**
- 6: gb_pat:**
- 7: gb_ph:**
- 8: gb_pl:**
- 9: gb_pr:**
- 10: gb_ro:**
- 11: gb_sts:**
- 12: gb_sy:**
- 13: gb_un:**
- 14: gb_vi:**
- 15: em_ba:**
- 16: em_fun:**
- 17: em_hum:**
- 18: em_in:**
- 19: em_mu:**
- 20: em_om:**
- 21: em_or:**
- 22: em_ov:**
- 23: em_pat:**
- 24: em_ph:**
- 25: em_pl:**
- 26: em_ro:**
- 27: em_sts:**
- 28: em_un:**
- 29: em_vi:**
- 30: em_htg_hum:**
- 31: em_htg_inv:**
- 32: em_htg_other:**
- 33: em_htg_mus:**
- 34: em_htg_pln:**
- 35: em_htg_rod:**
- 36: em_htg_mam:**
- 37: em_htg_vrt:**
- 38: em_sy:**
- 39: em_htgo_hum:**
- 40: em_htgo_mus:**
- 41: em_htgo_other:**

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	43.2	10.7	372	4	S75319	S75319 12S rRNA (n
2	41.6	10.3	239	11	G37943	G37943 C4M29 Plasm
3	39.8	9.8	375	11	BV014350	BV014350 S208P6381
C 4	38	9.4	255	8	AF351282	AF351282 Goseypum
C 5	37.2	9.2	283	6	BD211954	BD211954 Novel hum
6	37	9.1	263	11	CNS06EHA	AL395108 T3 end of
7	35.6	9.0	404	11	HUMUT5151	L17710 Human STS U
8	35.6	8.8	270	8	AY198661	AY198661 Arabidops
9	34.8	8.6	383	11	HUMUT5116	L18521 Human chrom
10	34.8	8.6	399	6	AR417479	AR417479 Sequence
11	34.8	8.6	399	6	BD113032	BD113032 EST and e
C 12	34.6	8.5	277	8	FSY566324	AJ566324 Ficus syc
C 13	34.4	8.5	372	9	BC015361	BC015361 Homo sapi
14	34.4	8.5	397	11	G47173	G47173 215801.1 Ze
15	34.2	8.4	353	11	CNS06EJ3	AL395173 T7 end of
16	34.2	8.4	401	3	AF543047	AF543047 Mesobuchu
17	34	8.4	230	11	G50943	G50943 SHGC-78669
C 18	33.8	8.4	296	10	AY040769S1	AY040769 Mus muscu
19	33.8	8.3	392	3	PFAHMG	L31630 Plasmodium
20	33.8	8.3	398	17	AF116644	AF116644 Homo sapi
21	33.6	8.3	249	6	AR413685	AR413685 Sequence
22	33.6	8.3	249	6	BD109238	BD109238 EST and e
23	33.6	8.3	385	3	PFAHMGH5	X94294 P.falciparu
24	33.4	8.2	240	6	AR413687	AR413687 Sequence
25	33.4	8.2	240	6	BD109240	BD109240 EST and e
C 26	33.2	8.2	190	9	AY376246	AY376246 Homo sapi
27	33.2	8.2	264	11	HUMUT5168B	L31078 Human STS U
28	33.2	8.2	288	11	HUMUT5193B	L31114 Human STS U
C 29	33.2	8.2	357	5	ASE276787	AJ276787 Acroceph
30	33	8.1	367	11	HUMUT8147	L30234 Human STS U
31	33	8.1	400	6	AR357935	AR357935 Sequence
32	32.8	8.1	294	3	AF201952	AF201952 Plasmodi
C 33	32.6	8.0	269	6	AX185675	AX185675 Sequence
34	32.4	8.0	310	11	BV077053	BV077053 S210P6110
35	32.4	8.0	352	3	AY054826	AY054826 Plasmodi
36	32.4	8.0	352	3	AY054852	AY054852 Plasmodi
C 37	32.4	8.0	381	11	BV095162	BV095162 RPAMSEQO
38	32.2	8.0	120	6	AX353334	AX353334 Sequence
39	32.2	8.0	333	6	AX363358	AX363358 Sequence
C 40	32	7.9	231	14	D67094	D67094 Hepatitis v
41	32	7.9	273	9	HUMD8321	L12269 Homo sapien
42	32	7.9	347	3	PFAHMG	M63272 P.falciparu
43	31.8	7.9	261	8	AF310859	AF310859 Betula pe
44	31.8	7.9	302	10	RNU00763	U00763 Rattus norv
C 45	31.8	7.9	384	6	BD244367	BD244367 Determina

ALIGNMENTS

RESULT 1

S75319

LOCUS

DEFINITION

S75319 372 bp DNA linear NAM 27-MAY-1995

12S rRNA (nuclear clone pNunt.1, tandem repeat macrosatellite)

[Felis catus=domestic cats, lymphocytes, fibroblasts, Genomic, 372

nt].

ACCESSION S75319

VERSION S75319.1

KEYWORDS GI:833954

SOURCE Felis catus (cat)

ORGANISM Felis catus

REFERENCE 1 (bases 1 to 372)

AUTHORS Lopez,J.V., Yuhki,N., Masuda,R., Modi,W. and O'Brien,S.J.

TITLE Numt, a recent transfer and tandem amplification of mitochondrial DNA to the nuclear genome of the domestic cat
 J. Mol. Evol. 39 (2), 174-190 (1994)
 MEDLINE 95018295
 PUBMED 7932781
 REMARK GenBank staff at the National Library of Medicine created this entry [NCBI gisbq 160474] from the original journal article. This sequence comes from Fig. 6B.
 Map location: D2.

FEATURES Location/Qualifiers
 source 1..372

gene /organism="Felis catus"
 /mol_type="genomic DNA"
 /db_xref="taxon:9685"
 1..372
 /gene="12S rRNA"

ORIGIN

Query Match 10.7%; Score 43.2; DB 4; Length 372;
 Best Local Similarity 57.4%; Pred. No. 1.8;
 Matches 78; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
 QY 95 AAATGCTACATTTGAAGAAACCTTCAAGATCAAGATAAGAAAGCCCAAGAGTTTCAT 154
 Db 170 ATATACCGGCATCTTCAGCAACCCCTAAGAGGAGAAAGATGACCAAGTATCTTAC 229
 QY 155 CCACCTTCTAATCAGGAACGAGAAATGGCAGTGGTTCTGAAGAGTGTGCTACACTGTCA 214
 Db 230 CCAAAAAAGTTAGGTCAAGGTGTAGCTCATGAGATGGGAGCAATGGCTACACTTCT 289
 QY 215 TTAATCAGATCCCCA 230
 Db 290 AAAATAGTAATACCCA 305

RESULT 2

G37943 239 bp DNA linear STS 26-FEB-1999
 LOCUS C4M29 Plasmodium falciparum haploid Plasmodium falciparum STS
 DEFINITION genomic, sequence tagged site.
 ACCESSION G37943
 VERSION G37943.2 GI:4337400
 KEYWORDS STS.
 SOURCE Plasmodium falciparum (malaria parasite P. falciparum)
 ORGANISM Plasmodium falciparum
 Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 1 (bases 1 to 239)
 Su, X.-z., Ferdig, M.T. and Wellens, T.E.
 Integrated genetic and physical maps of Plasmodium falciparum
 Unpublished (1998)
 TITLE On Mar 6, 1999 this sequence version replaced gi:3056687.
 JOURNAL COMMENT

Contact: Thomas E. Wellens

LFD

NIAID

Bldg. 4, Room 126, NIH Campus, Bethesda, MD 20892-0425

Tel: (301) 496-4021

Fax: (301) 402-0079

Email: tew@helix.nih.gov

Primer A: TCTTACACTTTTGTCCATCAT

Primer B: CATTTGGTACTGTAGTGATC

STS size: 239

PCR Profile:

Initial heat: 2 min at 94°C

Cycles: 30 cycles, each having steps 20sec at 94°C;

10sec at 45°C; 10sec at 40°C; 30sec at 60°C

Machine: Perkin Elmer GeneAmp 9600

Protocol:

Template: 1-10 ng

Primer: 5 pM each

dNTPs: 0.3 uM of 10 mM stock solution

Taq Polymerase: 0.5 unit

Total Volume: 15 uL

Buffer:

10 mM Tris HCl, pH 8.3
 1.5 mM MgCl2
 50 mM KCl.

FEATURES

Location/Qualifiers
 source 1..239
 /organism="Plasmodium falciparum"
 /mol_type="genomic DNA"
 /strain="HB3xDd2"
 /db_xref="taxon:5833"
 /map="4"

/clone_lib="Plasmodium falciparum haploid"
 /dev_stage="erythrocytic haploid stage"
 1..239

primer_bind 1..20

primer_bind complement (221..239)

ORIGIN

Query Match 10.3%; Score 41.6; DB 11; Length 239;
 Best Local Similarity 53.8%; Pred. No. 5;
 Matches 86; Conservative 0; Mismatches 74; Indels 0; Gaps 0;
 QY 41 AGAATCAAGAGAGCCCAAGAAAGGAAACCCAGATGAGGAAAGAAACCGCAGGAATCA 100
 Db 56 AAATAAAGAAAGAAATAAATAAAGAAAGAAATAAATAAAGAAAGAAATAAATAA 115
 QY 101 CTACATTTGAAAGAAACCTTCAAGATCAAGATAAGAAAGCCCAAGAGTTTTCATCCACTT 160
 Db 116 GAAAAATAAAGAAAGAAATAAATAAAGAAAGAAATAAATAAAGAAAGAAATAA 175
 QY 161 CTAATCAGGAACGAGAAATGGCAGTGGTTCTCAAGAACT 200
 Db 176 ATAAAAAGAAAGAAAGAAAGGGGGGTAGGGGAAGAAGT 215

RESULT 3

BV014350 375 bp DNA linear STS 30-MAY-2003
 LOCUS S208F6381FP2.T0 129S1/SvimJ Mus musculus STS genomic, sequence
 DEFINITION tagged site.
 ACCESSION BV014350
 VERSION BV014350.1 GI:31098245
 KEYWORDS STS.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 375)
 Wade, C.M., Kulbokas, E.J. III, Kirby, A.W., Zody, M.C., Mullikin, J.C.,
 Lander, E.S., Lindblad-Toh, K. and Daly, M.J.
 The mosaic structure of variation in the laboratory mouse genome
 Nature 420 (6915), 574-578 (2002)
 22354684
 12466852

COMMENT

Contact: Kerstin Lindblad-Toh
 Whitehead Institute for Biomedical Research, Center for Genome
 Research
 320 Charles Street, Cambridge, MA 02141, USA
 Tel: 6172521477
 Fax: 6172580903
 Email: kersli@genome.wi.mit.edu
 Primer A: None
 Primer B: None
 STS size: 375
 Protocol:
 WGS-discovery: Paired-end low-coverage whole genome shotgun reads
 were generated from 129S1/SvimJ, C3H/HeJ, and BALB/cByJ. The WGS
 reads were placed uniquely on the MGSCV3 C57BL/6J assembly and SNP
 detection was carried out by SSAHA-SNP. 225,000 reads were
 annotated
 as STSs and 81,000 SNPs were annotated with alleles from C57BL/6J

KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 270)
TITLE May, B.P., Smorowski, J., Arroyo, J.-M., Vaughn, M.W., Shen, R.,
 McCombie, W.R. and Martienssen, R.A.
JOURNAL Arabidopsis genomic sequences flanking Ds enhancer and gene traps
 in transgenic lines
REFERENCE Unpublished
AUTHORS 2 (bases 1 to 270)
TITLE May, B.P., Smorowski, J., Arroyo, J.-M., Vaughn, M.W., Shen, R.,
 McCombie, W.R. and Martienssen, R.A.
JOURNAL Direct Submission
COMMENT Submitted (17-DEC-2002) Cold Spring Harbor Laboratory, 1 Bungtown
 Road, Cold Spring Harbor, NY 11724, USA
FEATURES <http://genetrap.cshl.org>
 Location/Qualifiers
 source
 1..270
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /db_xref="taxon:3702"
 /cldnes="ET2091.Ds3.02.07.00.b.270"
 /ecotype="Landsberg"
 /note="transgenic line ET2091"
ORIGIN
 Query Match 8.8%; Score 35.6; DB 8; Length 270;
 Best Local Similarity 51.2%; Pred. No. 1.7e+02;
 Matches 83; Conservative 0; Mismatches 79; Indels 0; Gaps 0;
 QY 19 CGAAATCTGTTCTGCTGGGAGAGATCAAAAGAGCCGCAAGAGAAAGCAACCCAGATGAG 78
 Db 58 CAAAATGTAGTTCCTTTAAACAAGATCATATCGTTAAACAATCCAAAGCCCAAGAT 117
 QY 79 GAAAGAAACCGCAGGAAATGACTACATTTCAAGAAATCTTCAAGATCAAGATAGAAA 138
 Db 118 CGAAGAAACCAAGAGATGAAACACTATCTTTAAAGCTTATCGAGTGCATCATCA 177
 QY 139 AGCCAAGAGTTTCATCCATCTTCAATCAGGAAACGAGAAAT 180
 Db 178 AGCAAAACCCCTACAGAAATTCGAAAGATGAAAGAGGAGAAAT 219
RESULT 9
HUMUT5116 383 bp DNA linear STS 26-JUL-1993
LOCUS Human chromosome 22 STS UT5116, sequence tagged site.
DEFINITION L18521
ACCESSION L18521.1 GI:308187
VERSION STS: PCR primer; STS sequence; microsatellite marker;
KEYWORDS microsatellite repeat; repeat polymorphism; sequence tagged site;
 tetranucleotide repeat.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 383)
 Garken, S.C., Matsunami, N., Lawrence, E., Carlson, M., Moore, M.,
 Ballard, L., Mellis, R., Robertson, M., Bradley, P., Eisner, T.,
 Tingey, A., Rodriguez, P., Albertsen, H., Lalouel, J.-M. and White, R.
TITLE Genetic and physical mapping of simple sequence repeat containing
 sequence tagged sites from the human genome
JOURNAL Unpublished (1993)
COMMENT Original source text: Homo sapiens DNA.
 Submitted by: Utah Center for Human Genome Research University of
 Utah, Dept. of Human Genetics
 2160 Eccles Institute of Human Genetics
 Salt Lake City, UT 84112
 e-mail: sts@corona.med.utah.edu
 Primer A: TGCCACTGTACTCCAGCTT

Primer B: ARGATTATTGGCACCAAAACC
32P-label: A Primer
PCR Profile:
 Initial Denaturation: 94C 300sec
 PCR Cycles: 5
 Denaturation: 94C 10sec
 Annealing: 58C 10sec
 Extension: 72C 20sec
 Mg++: 1mM
 Gel: Acrylamide 7%, Formamide 32%, Urea 34%
Alleles: 4
FEATURES Location/Qualifiers
 source
 1..383
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /map="22"
 43..337
 /standard_name="STS UT5116"
 primer_bind 43..61
 primer_bind complement (317..337)
ORIGIN
 Query Match 8.6%; Score 34.8; DB 11; Length 383;
 Best Local Similarity 52.0%; Pred. No. 2.6e+02;
 Matches 78; Conservative 0; Mismatches 72; Indels 0; Gaps 0;
 QY 36 GGGAGAGATCAAAAGAGCCCAAGAGGAAACCCAGATGAGGAAAGAAACCGCAGGA 95
 Db 122 CGAAGGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 181
 QY 96 AATGACTCATTTTGAAGAGAAACCTTCAAGATCAAGATAGAAAGAAAGAAAGAAAGAAAG 155
 Db 182 AAAGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 241
 QY 156 CACTTCTTAATCAGGAAACGAGAAATGGCAG 185
 Db 242 GCGAGGAGGTGAGAGAGCGAGGAGGAAG 271
RESULT 10
AR417479 399 bp DNA linear PAT 18-DEC-2003
LOCUS AR417479
DEFINITION Sequence 8976 from patent US 6639063.
ACCESSION AR417479
VERSION AR417479.1 GI:40172589
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 399)
AUTHORS Edwards, J.-B.D.M., Jobert, S. and Giordano, J.-Y.
TITLE EST's and encoded human proteins
JOURNAL Patent: US 6639063-A 8976 28-OCT-2003;
FEATURES Location/Qualifiers
 source
 1..399
 /organism="unknown"
 /mol_type="genomic DNA"
ORIGIN
 Query Match 8.6%; Score 34.8; DB 6; Length 399;
 Best Local Similarity 12.4%; Pred. No. 2.6e+02;
 Matches 30; Conservative 110; Mismatches 102; Indels 0; Gaps 0;
 QY 3 GGGAAATATCTCTCTCGGAAACTCAGTTGCTGTGGAGAGAAATCAAAAGAGACCCCAAGAA 62
 Db 42 KCRERRRRRRMMWKKSCWKKSKSWRSWGMWTKMKRGGAASWAGYNSWWTYTRRWY 101
 QY 63 AGGAAACCCAGATGAGGAAAGAAAGAAACCGCAGGAATGACTACATTGGAAGAAACCTCA 122
 Db 102 YFYRKACTKWRAGMWGKAWYAKWYNAWRRTAMKYNAWMMKSKSRMRERRAWY 161
 QY 123 AGATCAAGATAAGAAAGCAAGAGAGTTTCATCTCACTTCTTAATCAGGAAAGAGAAATGG 182

```

162 WMYTMMARTWGWREASCYRGAYMASAGWYWMYMMRRKWMYSAGWSMKWTRRCAS 221
QY 183 CAGTGGTTCTCAGGAAGTGCTACACTGTCTAATACATCCCCCATCAGAGATCCTC 242
Db 222 YSCWSSYCWGAKMMYWKTSRWSYSSYRCKYRRSCCWSMSYWKYYSRWSYWCASC 281
QY 243 CC 244
Db 282 YY 283

RESULT 11
LOCUS BD113032 399 bp DNA linear PAT 18-SEP-2002
DEFINITION EST and encoded human protein.
ACCESSION BD113032
VERSION BD113032.1 GI:320207850
KEYWORDS JP 2002010789-A/5109
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 399)
AUTHORS Edwards, J.B.D.M., Jobert, S. and Giordano, J.E.
TITLE EST and encoded human protein
JOURNAL Patent: JP 2002010789-A 5109 15-JAN-2002;
GENSET CORP
OS Homo sapiens (human)
PN JP 2002010789-A/5109
PD 15-JAN-2002
PF 07-AUG-2000 JP 2000280989
PR 05-AUG-1999 US 60/147499
PI JEAN BAPTISTE DUMAS MILNE EDWARDS, SEVELIN JOBERT, JEAN EVE PI
GIORDANO
PC C12N15/09, C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, PC
C12N1/21,
PC C12N5/10, C12P21/02, C12P21/08, C12Q1/68, C12N15/00, C12N5/00, PC
C12N15/00
CC EST and encoded human protein
FH Key
FT source
FT 1.399
FT Location/Qualifiers
source
1.399
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 8.6%; Score 34.8; DB 6; Length 399;
Best Local Similarity 12.4%; Pred. NO. 2.6e+02;
Matches 30; Conservative 110; Mismatches 102; Indels 0; Gaps 0;

QY 3 GGGAAATATCTCCGCGAAACTCAGTTGCTGGGAGAGATCAAAAGAGCCCAAGAA 62
Db 42 KRRRRRAWWKSKWKKSKSWRGWMTTKKRGKGAASWAGYMSWMTYTRRWY 101
QY 63 AGGAACCCAGATGAGGAAGAAAGCCGAGGAATGACTATTTGAAGAAACTTCA 122
Db 102 YVYRKCACTKWAAGWGWKAGWAWAYAKWYMWARTAMKYWAMKSKSRMRRAWYA 161
QY 123 AGATCAAGTAAGAAAGCCAGAGATTTCATCCACTTCTAATCAGGAACAGAGATGG 182
Db 162 WMYTMMARTWGWREASCTRGYMASAGWYWMYMMRRKWMYSAGWSMKWTRRCAS 221
QY 183 CAGTGGTTCTCAGGAAGTGCTACACTGTCTAATACATCCCCCATCAGAGATCCTC 242
Db 222 YSCWSSYCWGAKMMYWKTSRWSYSSYRCKYRRSCCWSMSYWKYYSRWSYWCASC 281
QY 243 CC 244
Db 282 YY 283
```

```

RESULT 12
LOCUS FSV566324/c 277 bp DNA linear PLN 30-JUN-2003
DEFINITION Ficus sycamoros microsatellite DNA, clone FIGT09D12.
ACCESSION AJ566324
VERSION AJ566324.1 GI:32399995
KEYWORDS microsatellite; repetitive DNA.
SOURCE Ficus sycamoros
ORGANISM Ficus sycamoros
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Rosales; Moraceae; Ficus.
REFERENCE 1
AUTHORS Dawson, D.A., Ahmed, S., Compton, S.G., Gilmartin, P.M. and Burke, T.
TITLE Isolation of microsatellite loci in the fig, Ficus sycamoros
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 277)
AUTHORS Dawson, D.A.
TITLE Direct Submission
JOURNAL Submitted (28-MAY-2003) Dawson D.A., Department of Animal and Plant
Sciences, University of Sheffield, Alfred Denny Building, Western
Bank, Sheffield, S10 2TN, UNITED KINGDOM
FEATURES
source
1..277
/organism="Ficus sycamoros"
/mol_type="genomic DNA"
/db_xref="taxon:182129"
/clone="FIGT09D12"
87..190
satellite
1.277
/feature="microsatellite"
rpt_unit="(T)6(C)1(T)6(G)1(T)7(A)2(TTTC)1(TTTC)1
(TTTC)5(TTTC)7(C)1(T)18"
/rpt_type="TANDEX"

ORIGIN
Query Match 8.5%; Score 34.6; DB 8; Length 277;
Best Local Similarity 58.1%; Pred. NO. 3.1e+02;
Matches 61; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 39 AGGATCAAAAGAGCCCAAGAGAGGAAACCAGATGAGGAAGAAACGCGAGAAAT 98
Db 188 AAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 129
QY 99 GACTACATTGAAAGAAAACTTCAAGATCAAGATCAAGATCAAGAGAAAGCCCA 143
Db 128 GAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 84

RESULT 13
LOCUS BC015361 372 bp mRNA
DEFINITION Homo sapiens, clone IMAGE:3886095, mRNA.
ACCESSION BC015361
VERSION BC015361.1 GI:15929880
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 372)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (01-OCT-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: DCID/DPF/Gazdar
cDNA Library Preparation: Life Technologies, Inc.
```

RESULT 15

CNS06UEJ3	353 bp	DNA linear	STS 10-JAN-2001
CNS06UEJ3	T7 end of clone ARORA022A02 of library AR0AA from strain CBS 732 of Zygosaccharomyces rouxii, sequence tagged site.		
DEFINITION			
ACCESSION	AL395173		
VERSION	AL395173.1		
KEYWORDS	GI:12146610		
SOURCE	STS		
ORGANISM	Zygosaccharomyces rouxii		
	Zygosaccharomyces rouxii		
	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;		
	Saccharomycetales; Saccharomycetaceae; Zygosaccharomycetes.		
REFERENCE	1 (bases 1 to 353)		
AUTHORS	Soucié,J.L., Aigle,M., Artiguenave,F., Blandin,G., Bolotin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S., de-Montigny,J.M., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,		

	Malpertuy,A., Neuveglise,C., Ozier-Kalogeropoulos,O., Potier,S., Saurin,W., Tekala,F., Toffano-Nioche,C., Wesolowski-Jouvet,M., Winkler,P. and Weissenbach,J. Genomic exploration of the hemiascomycetous yeasts: 1. A set of yeast species for molecular evolution studies FEMS Lett. 487 (1), 3-12 (2000)
JOURNAL	20584711
MEDLINE	11152876
PubMed	11152876
REFERENCE	2 (bases 1 to 353)
AUTHORS	de Monigny,J., Straub,M., Potier,S., Tekala,F., Dujon,B., Winkler,P., Artiguenave,F. and Souciet,J. Genomic exploration of the hemiascomycetous yeasts: 8. Zygosaccharomyces rouxii FEMS Lett. 487 (1), 52-55 (2000)
JOURNAL	20584718
MEDLINE	11152883
PubMed	11152883
REFERENCE	3 (bases 1 to 353)
AUTHORS	Genoscope.
REFERENCE	Direct Submission
JOURNAL	Submitted (06-SEP-2000) Genoscope - Centre National de Sequencage, 2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail : this@genoscope.cns.fr - Web : www.genoscope.cns.fr) This STS is part of a random genomic sequencing program of thirteen yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii, Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila, Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to 5 kb were prepared and both extremities were sequenced. See keywords for description of this sequence and for the sequence of the other extremity of this insert.
FEATURES	Location/Qualifiers 1..353 /organism="Zygosaccharomyces rouxii" /mol_type="genomic DNA" /strain="CBS 732" /db_xref="taxon:4956" /clone="AR0A022A02" /clone_lib="AR0A" /note="end : T7" 1..353
STS	
ORIGIN	
Query Match	8.4%; Score 34.2; DB 11; Length 353;
Best Local Similarity	54.3%; Pred. NO. 3.8e+02;
Matches	69; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
QY	21 AAAAAGCTCAGTTCCTGGGAGAGAAATCAAAAGAGAGCCCAAGAGGAAAGCCGATGAGGA 80
Db	153 AAAACAAAAAACAATTAATAAAAAAAAAAAAAAAAAAGAGAAAAACCAAGAAAA 212
QY	81 AGAGAAACGGCGGAATGACTACATTTGAAGACAAACCTTCAGATCAAGATAAGAAAAAG 140
Db	213 AAAAAATTCAAAAAAATAAATAAGGAAAAAATAAATAAGAAATTAATAAAAAAAAAAAAA 272
QY	141 CCAAGAA 147
Db	273 AAAAAAA 279

Search completed: April 23, 2004, 03:08:42
Job time : 2081 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 22, 2004, 18:29:17 ; Search time 2022 Seconds
(without alignments)

5981.297 Million cell updates/sec

Title: US-10-005-907-1_COPY_25_429

Perfect score: 405
Sequence: 1 atgggaattatctctgcg.....atgaagtgtgtttccacac 405

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 16189338

Minimum DB seq length: 0
Maximum DB seq length: 405

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- EST.*
- 1: em_estba.*
 - 2: em_estum.*
 - 3: em_estin.*
 - 4: em_estmd.*
 - 5: em_estov.*
 - 6: em_estpl.*
 - 7: em_estro.*
 - 8: em_hci.*
 - 9: gb_est1.*
 - 10: gb_est2.*
 - 11: gb_hci.*
 - 12: gb_est3.*
 - 13: gb_est4.*
 - 14: gb_est5.*
 - 15: em_estfun.*
 - 16: em_estom.*
 - 17: em_gss_hum.*
 - 18: em_gss_inv.*
 - 19: em_gss_pln.*
 - 20: em_gss_vrt.*
 - 21: em_gss_fun.*
 - 22: em_gss_mam.*
 - 23: em_gss_mus.*
 - 24: em_gss_pro.*
 - 25: em_gss_rod.*
 - 26: em_gss_phg.*
 - 27: em_gss_vrl.*
 - 28: gb_gss1.*
 - 29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54.8	16.0	288	10 AW480906	AW480906 33516 MAR
2	46.8	11.6	353	29 AY404587	AY404587 Pan trogl
3	46.6	11.5	259	29 CE192340	CE192340 tigr-gss-
4	44	10.9	203	29 CE091328	CE091328 tigr-gss-

5	43.4	10.7	404	13 BX367419	BX367419
6	42	10.4	345	14 CF351413	CF351413 lab89e12.
7	41.8	10.3	160	28 AZ273050	AZ273050 RPCI-23-8
8	41.8	10.3	274	29 CE329306	CE329306 tigr-gss-
9	41.2	10.2	357	28 AZ419893	AZ419893 1M0196A08
c 10	40.6	10.0	276	28 AQ645726	AQ645726 RPCI93-EC
11	40.6	10.0	359	13 BX460850	BX460850 BX460850
12	40.4	10.0	244	9 AL725744	AL725744 AL725744
c 13	39.8	9.8	360	29 CNS020HP	CNS020HP
14	39.8	9.8	261	29 CNS0232W	CNS0232W
15	39.6	9.8	372	9 AU039357	AU039357 AU039357
c 16	39.4	9.7	261	29 CNS006BK	CNS006BK
c 17	39.4	9.7	364	29 CE343646	CE343646 tigr-gss-
18	39.4	9.7	395	28 AZ937971	AZ937971 2M0198B23
19	39.4	9.7	399	14 CD677582	CD677582 hol19b06.y
20	39.2	9.7	356	14 CF379377	CF379377 lac85e12.
21	39	9.6	358	14 CF803201	CF803201 rd95d09.y
22	39	9.6	401	9 AL134217	AL134217 DKFP547B
23	38.8	9.6	210	9 AA225339	AA225339 nc23h08.s
c 24	38.8	9.6	364	29 CE462660	CE462660 tigr-gss-
25	38.6	9.5	252	12 EG981526	EG981526 MR3-CH014
c 26	38.6	9.5	268	12 BI493055	BI493055 d132908.w
27	38.6	9.5	382	12 EG497771	EG497771 602543034
28	38.6	9.5	389	10 AW786169	AW786169 119070 MA
29	38.6	9.5	400	12 EG382689	EG382689 299593 MA
30	38.4	9.5	244	29 CE825890	CE825890 tigr-gss-
31	38.4	9.5	279	12 BI671444	BI671444 fs48d07.y
32	38.4	9.5	292	14 CD422569	CD422569 laa82h05.
33	38.4	9.5	293	12 EM155090	EM155090 fv93d10.y
c 34	38.4	9.5	320	13 BX361849	BX361849 BX361849
35	38.4	9.5	355	29 CE797846	CE797846 tigr-gss-
36	38.2	9.4	186	13 C93279	C93279 C93279 Dict
37	38.2	9.4	313	29 CNS00CZT	CNS00CZT
c 38	38.2	9.4	327	29 CE476461	CE476461 Drosophil
39	38.2	9.4	357	29 CE302878	CE302878 tigr-gss-
c 40	38	9.4	131	14 CF330236	CF330236 NACL--05-
41	38	9.4	210	12 BM873785	BM873785 laa03a07.
42	38	9.4	250	14 CB289843	CB289843 kb53b05.y
43	38	9.4	278	14 CF544822	CF544822 lad83b08.
c 44	38	9.4	304	10 BB291389	BB291389 BB291389
45	38	9.4	383	14 CF804314	CF804314 lad70c09.

ALIGNMENTS

RESULT 1
AW480906 288 bp mRNA linear EST 09-JUL-2000
LOCUS AW480906
DEFINITION 33516 MARC 2PIG Sus scrofa cDNA 5', mRNA sequence.
ACCESSION AW480906
VERSION AW480906.1 GI:7050949
KEYWORDS EST.
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (bases 1 to 288)
AUTHORS Vallet,J., Wise,T., Rohrer,G.A., Pertea,G., Sultana,R.,
Fahrenkrug,S.C., Smith,T.P.L., Freking,B.A., Cho,J., White,J.,
Quackenbush,J. and Keele,J.W.
TITLE Porcine gene discovery by normalized cDNA-library sequencing and
EST cluster assembly
JOURNAL Mamm. Genome 13 (8), 475-478 (2002)
MEDLINE 22213789
COMMENT 12226715
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and trimmed with phred

v0.90904.e. Vector identified by cross_match with the -minscore 20 and -minmatch 12 options.

PCR Primers
 FORWARD: AGGAACAGCTATGACCAT
 BACKWARD: GTTATCCAGTCACGACG
 Plate: 19 row: F column: 21
 Seq primer: ATTAGTGACACATATAG.
 Location/Qualifiers

FEATURES

source

1..288

/organism="Sus scrofa"
 /mol_type="mRNA"
 /db_xref="taxon:9823"
 /tissue_type="pooled"
 /lab_host="DH10B"
 /clone_lib="MARC 2P1G"
 /note="vector: pcMV SPORT6; Site 1: NotI; Site 2: SalI;
 Library made from pooled tissue from testis, ovary,
 endometrium, hypothalamus, pituitary, and placenta."

ORIGIN

Query Match 16.0%; Score 64.8; DB 10; Length 288;
 Best Local Similarity 75.8%; Pred. No. 4e-05;
 Matches 94; Conservative 0; Mismatches 27; Indels 3; Gaps 1;
 QY 285 CACAAGGAAGTGAAGACCTTACAGAAAGTTCAGACACAGATATGCCCTTCTTAGGAC 344
 DB 2 CACAAGAGAGTGAAGACCTTACAGAAAGTTCAGACACAGATATGCCCTTCTTAGGAC 61
 QY 345 TTCGTAGT---AGGCTTCTTCTGACCCATGAGCATGATTAAGATTGTTTCC 401
 DB 62 GACTTACACTCAAGTTTTCTTCTATACCCCTGAGATGATTAAGATTGTTTCC 121
 QY 402 ACAC 405
 DB 122 TCAC 125

RESULT 2
 AY404587
 LOCUS
 DEFINITION Pan troglodytes GCET2 gene, VIRTUAL TRANSCRIPT, partial sequence,
 genomic survey sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE

ORGANISM
 Pan troglodytes (chimpanzee)

REFERENCE
 AUTHORS
 Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
 Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
 Adams,M.D. and Cargill,M.

TITLE
 Inferring nonneutral evolution from human-chimp-mouse orthologous

gene trios

JOURNAL
 PUBLISHED
 REFERENCE

2 (bases 1 to 353)
 Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
 Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
 Adams,M.D. and Cargill,M.

Direct Submission

TITLE
 Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA

COMMENT
 This sequence was made by sequencing genomic exons and ordering
 them based on alignment.

FEATURES
 Location/Qualifiers

1..353
 /organism="pan troglodytes"
 /mol_type="genomic DNA"
 /db_xref="taxon:9598"

gene

<1..>353
 /gene="GCET2"
 /locus_tag="HCM1940"

ORIGIN

Query Match 11.6%; Score 46.8; DB 29; Length 353;
 Best Local Similarity 47.7%; Pred. No. 0.77; 113; Indels 3; Gaps 1;
 Matches 106; Conservative 0; Mismatches 54; Indels 0; Gaps 0;
 QY 114 AAACCTTCAAGATCAAGATAAGAAAGCCAAAGAGTTTTCATCCACTTCTTAATCAGAAAA 173
 DB 131 AAAGANNNAAGATTCCCAAAACGAAATGAAGGAATGTCACTACCTCCATCCAGNNNN 190
 QY 174 CAGAAATGGCAGTGGTCTTGAAGAAGTGTGTACACTGTCTTAATATCA---TCCCCCA 230
 DB 191 NNNNNNNNNNNNNNNNNNNAGNAGCTGTCTANANCNTCAATNNNCGGTTCTCTG 250
 QY 231 TCAGAGATCCTCCCTGAGCTCCAATGATGATGGCTATGAGAACATTGACTCCCTCACAAG 290
 DB 251 NACRAGGCCATCAGGGAACCTCTGCTGAAGACTACTATGAGATGTTCCTCGAAGCTGA 310
 QY 291 GAAATGAGACAGTTTGAAGAAAGTTCAGACACAGATATGC 332
 DB 311 GAGACCCAGAGAGCCCTTTGGGAGGAACCTGAGACTGAGTATTC 352

RESULT 3

CE192340/c

CE192340/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

Contact: Kirkness EF

The Institute for Genomic Research

Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,

Rockville, MD 20850, USA

Tel: 301-838-0200

Fax: 301-838-0208

Email: ekirknes@tigr.org

Class: shotgun.

Location/Qualifiers

1..259

/organism="Canis familiaris"

/mol_type="genomic DNA"

/strain="Standard Poodle"

/db_xref="taxon:9615"

/clone_lib="Dog Library"

/note="Site 1: BstXI; Libraries were prepared from

peripheral blood"

ORIGIN

Query Match

Best Local Similarity

Matches

79; Conservative

11.5%; Score 46.6; DB 29; Length 259;

Pred. No. 0.84;

Mismatches 54; Indels 0; Gaps 0;

QY 47 AAAGAAGCCCAAGAAAGAAACCCAGATGAGAAAGAAACCGGAGGAATGACTACAT 106

DB 166 AAAGAAGCCCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 107

```

QY 107 TTGAAGAAACTTCAAGATCAAGATAAGAAAGCCAGAGAGTTTCATCCATCTTCTTAATC 166
Db 106 AAGAAAGAAAGAGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 47
QY 167 AGGAAACAGGAA 179
Db 46 AAAGAAAGAAAGAA 34

RESULT 4
CE091328 203 bp DNA linear GSS 24-SEP-2003
LOCUS tigr-gss-dog-17000359492028 Dog Library canis familiaris genomic,
DEFINITION genomic survey sequence.
ACCESSION CE091328
VERSION CE091328.1 GI:35158174
KEYWORDS GSS.
SOURCE Canis familiaris (dog)
ORGANISM Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE 1 (bases 1 to 203)
AUTHORS Kirkness,E.F., Bafna,V., Halpern,A.L., Levy,S., Remington,K.,
Rusch,D.B., Delcher,A.L., Pop,M., Wang,W., Fraser,C.M. and
Venter,J.C.
TITLE The dog genome: survey sequencing and comparative analysis
JOURNAL Science 301 (5641), 1898-1903 (2003)
MEDLINE 22875432
PUBMED 14512627
COMMENT Contact: Kirkness EF
The Institute for Genomic Research
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
Rockville, MD 20850, USA
Tel: 301-838-0200
Fax: 301-838-0208
Email: ekirknes@tigr.org
Class: shotgun.
FEATURES
source
1..203
Location/Qualifiers
/organism="Canis familiaris"
/mol_type="genomic DNA"
/strain="Standard Poodle"
/db_xref="taxon:9615"
/clone_lib="Dog Library"
/note="Site 1: BstXI; Libraries were prepared from
peripheral blood"
ORIGIN
Query Match 10.9%; Score 44; DB 29; Length 203;
Best Local Similarity 55.1%; Pred. No. 3.4;
Matches 86; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 20 GAAAACTCAGTTGCTGGGAGAGATCAAAAGAGCCCAAGAGAAAGCCAGATCAGG 79
Db 20 GAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 79
QY 80 AAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 139
Db 80 AAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 139
QY 140 GCCAGAGAGTTTCATCCACTTCAATCAAGAGAAAGC 175
Db 140 GAAAGAAAGATGTTTATTCCTCTGGAAAGAAATG 175

RESULT 5
BX367419 404 bp mRNA linear EST 08-MAY-2003
LOCUS BX367419
DEFINITION Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
ACCESSION BX367419
VERSION BX367419.1 GI:30455633
KEYWORDS EST.

```

```

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 404)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 10480.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0AP0042G06NP1&cluster=10480.f. Contact :
Peng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0AP0042G06NP1.
FEATURES
source
1..404
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DJ009YM22"
/cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
/cell_line="JURKAT"
/clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT
10-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match 10.7%; Score 43.4; DB 13; Length 404;
Best Local Similarity 55.0%; Pred. No. 5;
Matches 71; Conservative 6; Mismatches 52; Indels 0; Gaps 0;

QY 19 CGAAAACTCAGTTGCTGGGAGAGATCAAAAGAGCCCAAGAGAAAGCCAGATCAG 78
Db 191 CCCGACTCAGATGCTGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 250
QY 79 GAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 138
Db 251 AAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 310
QY 139 AGCCAGAA 147
Db 311 AAAAGAGAA 319

RESULT 6
CF351413 345 bp mRNA linear EST 20-AUG-2003
LOCUS lab99el2.y1 SIEP Mus musculus CDNA 5', mRNA sequence.
DEFINITION CF351413
ACCESSION CF351413.1 GI:33954424
VERSION CF351413.1
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 345)
AUTHORS Tidwell,R., Clifton,S., Marra,M., Hillier,L., Pape,D., Martin,J.,
Wylie,T., Theising,B., Bowers,Y., Gibbons,M., Ritter,E., Bennet,J.,
Ronko,I., Tsagarashvili,R., Belaygorod,L., Grow,A., Maguire,L.,
Waterston,R. and Wilson,R.
TITLE WashU Stem cell EST Project
JOURNAL Unpublished (2002)
COMMENT Contact: Jeff Gordon and Mike Lovett
WashU, Human Genetics Division

```

Washington University School of Medicine
Library materials provided by: Stappenbeck Th.S. and Gordon J.I.
Library constructed by: Korshunova I. and Lovett M. DNA sequencing
by: Washington University Genome Sequencing Center For information
on obtaining a clone please contact: Rose Tidwell
(rtidwell@genetics.wustl.edu)
Putative full length read
vector to vector length is
Seq primer: -40RP from Gibco.
Location/Qualifiers
1. 345
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/tissue_type="small intestinal epithelial progenitors"
/lab_host="DMSalpha Ultra Max cells (Invitrogen)"
/clone_lib="sIEP"
/notes="vector: pAMP1; The library was synthesized with
modified SMART primers with dUTP at the end. After
treatment with UDG the cDNA was cloned in pAMP1 vector by
annealing."

FEATURES

source

ORIGIN

Query Match 10.4%; Score 42; DB 14; Length 345;
Best Local Similarity 52.9%; Pred. No. 11;
Matches 90; Conservative 0; Mismatches 80; Indels 0; Gaps 0;
QY 14 TCCTGCGAAACTCAGTTCCTGGAGAGATCAAAAGAGCCCAAGAGGAAACCCAG 73
Db 5 TGCTATTAACTCTACTAAGGGGGAAGAAAAAAGAAAAAAGAAAAAAGAAAA 64
QY 74 ATGAGGAAGAAAGCGGAGGAATGCTACTATTGAAAGAAACTTCAGATCAAGATA 133
Db 65 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 124
QY 134 AGAAAGCCAGAGAGTTCTATCCACTTCTATCAGGAAACGAGATGGC 183
Db 125 AAAAAAAGAAAAAGTCCCTCCCTTTTACCCCAAAAAAAGGCC 174

RESULT 7
LOCUS AZ273050 160 bp DNA linear GSS 26-JUL-2000
DEFINITION RPCI-23-85A5 TJC RPCI-23 Mus musculus genomic clone RPCI-23-85A5,
genomic survey sequence.
ACCESSION AZ273050
VERSION AZ273050.1 GI:9486590
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 160)
Zhao, S., Nieman, W., Feldblyum, T., Malek, J., Shatsman, S.,
Akinret, B., Levins, M., McGann, S., Teegaye, G., Geer, K., Krol, M., de
Jong, P. and Fraser, C.M.
Mouse BAC End Sequences from Library RPCI-23
Unpublished (1999)
Other GSSs: RPCI-23-85A5.TV
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPCI-23. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (<http://bacpac.med.buffalo.edu/orderingframe.htm>)
or from Resea ch Genetics (info@resgen.com). BAC end page:
http://www.tigr.org/tldb/bac/bac_ends/mouse/bac_end_intro.html
Plate: 85 row: A column: 5

Seq primer: SP6
Class: BAC ends.
Location/Qualifiers
1. 160
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-23-85A5"
/sex="Female"
/lab_host="DH10B"
/clone_lib="RPCI-23"
/note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1:
ECORI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or
brain genomic DNA was isolated and partially digested
with a combination of EcoRI and EcoRI Methyase. Size
selected DNA was cloned into the pBACe3.6 vector at the
ECORI sites. The ligation products were transformed into
DH10B electrocompetent cells (BRL Life Technologies)."

ORIGIN

Query Match 10.3%; Score 41.8; DB 28; Length 160;
Best Local Similarity 58.4%; Pred. No. 11;
Matches 73; Conservative 0; Mismatches 52; Indels 0; Gaps 0;
QY 22 AAATCAGTTCCTGGAGAGATCAAAAGAGCCCAAGAGGAAACCCAGATGAGAA 81
Db 36 AAATCAGTTCCTGGAGAGATCAAAAGAGCCCAAGAGGAAACCCAGATGAGAA 95
QY 82 AGAAAGCCAGAGGAAATGACTACTATTTGAAAGAAACTTCAAGATCAAGATAAGAAAAAGC 141
Db 96 AAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 155
QY 142 CAAGA 146
Db 156 AAGA 160

RESULT 8

CE329306 274 bp DNA linear GSS 26-SEP-2003
LOCUS tigr-gss-dog-17000333970436 Dog Library Canis familiaris genomic,
genomic survey sequence.
DEFINITION CE329306
ACCESSION CE329306.1 GI:36144095
VERSION CE329306
KEYWORDS GSS.
SOURCE Canis familiaris (dog)
ORGANISM Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
1 (bases 1 to 274)
Kirkness, E.F., Bafna, V., Halpern, A.L., Levy, S., Remington, K.,
Rusch, D.B., Delcher, A.L., Pop, M., Wang, W., Fraser, C.M. and
Venter, J.C.
The dog genome: survey sequencing and comparative analysis
Science 301 (5641), 1898-1903 (2003)
22875432
14512627
Contact: Kirkness EF
The Institute for Genomic Research
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
Rockville, MD 20850, USA
Tel: 301-838-0200
Fax: 301-838-0208
Email: ekirknes@tigr.org
Class: shotgun.

FEATURES

source

Location/Qualifiers
1. 274
/organism="Canis familiaris"
/mol_type="genomic DNA"
/strain="Standard Poodle"
/db_xref="taxon:9615"
/clone_lib="Dog Library"

/notes="Site 1: BstXI; Libraries were prepared from peripheral blood"

ORIGIN

Query Match 10.3%; Score 41.8; DB 29; Length 274;
Best Local Similarity 61.5%; Pred. No. 12;
Matches 67; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 39 AGAGATCAAAAGAGCCCAAGAAAGAAACCCAGATGAGAAAGAAACCGCAGGAAT 98
|||||
Db 18 AGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 77
|||||

QY 99 GACTACATTTGTAAGAAATCTTCAGATCAAGATCAAGAAAGAAAGCCAGAA 147
|||||
Db 78 AAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 126
|||||

RESULT 9
AZ419893

LOCUS 357 bp DNA linear GSS 03-OCT-2000

DEFINITION IM0196A08R Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0196A08 R, genomic survey sequence.

ACCESSION AZ419893
VERSION AZ419893.1 GI:10543906

KEYWORDS Mus musculus (house mouse)

SOURCE Mus musculus

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 357)

AUTHORS Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C., Riley, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A. and Wright, D., Weiss, R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

JOURNAL Unpublished (2000)

COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0196 row: A column: 08
Seq primer: CACACAGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 357.

FEATURES

Location/Qualifiers

1. .357

/organism="Mus musculus"

/mol_type="genomic DNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clones="UGC1M0196A08"

/sex="Male"

/lab_host="E. Coli strain XL10-Gold, TI-resistant, F-"

/clone_lib="Mouse 10kb plasmid UUGC1M library"

/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PWD42 (gi|4732114|gb|AF129072.1), a copy-number

inducible derivative of plasmid RI. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN

Query Match 10.2%; Score 41.2; DB 28; Length 357;
Best Local Similarity 60.9%; Pred. No. 17;
Matches 67; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 39 AGAGATCAAAAGAGCCCAAGAAAGAAACCCAGATGAGAAAGAAACCGCAGGAAT 98
|||||
Db 64 AAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 123
|||||

QY 99 GACTACATTTCAAGAAACTTCAAGATCAAGATCAAGAAAGAAAGCCCAAG 148
|||||
Db 124 AAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 173
|||||

RESULT 10

AQ645726/c

LOCUS 276 bp DNA linear GSS 08-JUL-1999

DEFINITION RPCI93-EcoRI-4J12.TV RPCI93-EcoRI Trypanosoma brucei genomic clone

ACCESSION AQ645726

VERSION AQ645726.1 GI:5122436

KEYWORDS Trypanosoma brucei

SOURCE Trypanosoma brucei

ORGANISM Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.

REFERENCE 1 (bases 1 to 276)

AUTHORS El-Sayed, N., Zhao, S., Zhao, H., Gill, S., Sub, E., Malek, J., Fujii, C., Gerrard, C., Leech, V., de Jong, P., Ullu, E., Melville, S., Doneison, J., Fraser, C. and Adams, M.

TITLE Use of BAC end sequences from Trypanosoma brucei GUTat 10.1 RPCI-93 Library for gene discovery and sequence-ready map construction

JOURNAL

COMMENT

Other GSSs: RPCI93-EcoRI-4J12.TJ
Contact: Najib M. El-Sayed
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208

Email: nelsayed@tigr.org

Clones and high density filters may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu). BAC end sequences search page: http://www.tigr.org/cdb/mbd/tbdb/.

Seq primer: T7

Class: BAC ends.

Location/Qualifiers

FEATURES

source

1. .276

/organism="Trypanosoma brucei"

/mol_type="genomic DNA"

/strain="TREU927/4 GUTat 10.1"

/db_xref="taxon:5691"

/clones="RPCI93-EcoRI-4J12"

/clone_lib="RPCI93-EcoRI"

/note="Vector: pBACe3.6; Site 1: Eco RI; Site 2: Eco RI; Constructed for The Institute for Genomic Research by Bohui Zhao in Pieter de Jong's laboratory (Roswell Park Cancer Institute, Buffalo, NY). Briefly, Trypanosoma brucei TREU927/4 GUTat 10.1 agarose embedded DNA was partially digested with a combination of Eco RI and Eco RI methylase (RPCI93-EcoRI segment) or Dpn II (RPCI93-DpnII segment). High molecular weight fragments were ligated in pBACe3.6 vector digested with Eco RI or Bam HI, respectively. The average insert size is 141 Kb. Total coverage (both segments): > 90 X the haploid non-minichromosomal genome."

ORIGIN

Query Match 10.0%; Score 40.6; DB 28; Length 276;
 Best Local Similarity 57.5%; Pred. No. 23;
 Matches 73; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 22 AAATCTAGTTGCTGGAGAGATCAAAAGAGCCCAAGAAAGCCAGATGAGGAA 81
 |||||
 Db 159 AAGCAGAGCTTCAGAGAAAGATGCAATGCAGCTGGAGACACAAACTGATGTGAA 100
 |||||

QY 82 AGAAACGGCAGGAATGACTTACATTTGAAGAAACTTCAAGATCAAGATAGAAAGC 141
 |||||
 Db 99 AGAAACCTGGTGCATTTAGACGCCAACAAACAAAGAGTCACATTCAGTGAAGAACCA 40
 |||||

QY 142 CAAGAAG 148
 |||||

Db 39 AACCAAG 33
 |||||

RESULT 11
 BX460850
 LOCUS
 DEFINITION BX460850 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
 CSODF019VB18 3-PRIME, mRNA sequence.

ACCESSION BX460850
 VERSION BX460850.1 GI:31027265
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 359)
 Li W.B., Gruber C., Jessee, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 5981.f For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 http://bin/cluster.cgi?seq=CSODF019DA09NP1&cluster=5981.f. Contact :
 Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CSODF019DA09NP1.
 Location/Qualifiers
 1..359
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CSODF019VB18"
 /tissue_type="FETAL BRAIN"
 /dev_stage="fetal"
 /clone_lib="Homo sapiens FETAL BRAIN"
 /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
 was primed with a NotI-oligo(dT) primer. Five prime end
 enriched, double-strand cDNA was digested with Not I and
 cloned into the Not I and EcoRV sites of the pCMVSPORT 6
 vector. Library was not normalized."

FEATURES

source

Query Match 10.0%; Score 40.6; DB 13; Length 359;
 Best Local Similarity 49.1%; Pred. No. 23;
 Matches 79; Conservative 9; Mismatches 73; Indels 0; Gaps 0;

QY 19 CGAAACTCAGTTGCTGGAGAGATCAAAAGAGCCCAAGAAAGCCAGATGAG 78
 |||||
 Db 129 CCAAAAACAACTCTGAGGAGAAAAGAAAAAAGAAAAAAGCAAAATCAGGAATR 188
 |||||

QY 79 GAAGAAACCGCAGGAATGACTCATTTGAAGAAAGAACTTCAGATCAAGATAAGAA 138
 |||||

ORIGIN

Query Match 10.0%; Score 40.4; DB 9; Length 244;
 Best Local Similarity 50.5%; Pred. No. 25;
 Matches 98; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 10 TATCTCTCGGAAACTCAGTTGCTGGAGAGATCAAAAGAGCCCAAGAAAGGAAAC 69
 |||||
 Db 51 TCTCTGATCGACACTGTTATTCGCTTATGAGATGAATAATAAATTAATTGAAAT.110
 |||||

QY 70 CCAGATGAGAAAGAAACCGCAGGAAATGACTATTCATCCACTTCTATTCAGGAAACGAGATGCGAGTGT 189
 |||||
 Db 111 TAAATGAAATCAAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 170
 |||||

QY 130 GATAAGAAAGCCAGAGACTTTTCATCCACTTCTATTCAGGAAACGAGATGCGAGTGT 189
 |||||
 Db 171 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGTCTCTGG 230
 |||||

QY 190 TCTGAAGAAAGTGTG 203
 |||||

Db 231 CCGGAACCAACCCCTG 244
 |||||

RESULT 13

CNS02OHP/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

CNS02OHP
 Tetraodon nigroviridis genome survey sequence PUC-Ori end of clone
 153J05 of library G from Tetraodon nigroviridis, genomic survey
 sequence.
 AL206710
 AL206710.1 GI:7865529
 GSS; genome survey sequence.

Db 189 AAAAAAAGAAAAAGAAAGAAATAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 248
 |||||

QY 139 AGCCAAAGAGTTTCATCCACTTCTAATCAGGAAACCGAGAA 179
 |||||

Db 249 AAACAAAGAAAAAAGAAAAAATAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 289
 |||||

RESULT 12

AL725744

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

ORIGIN

Query Match

Best Local Similarity

Matches

Conservative

Mismatches

Indels

Gaps

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

AL178961.1	GI:7817018
GSS:	genome survey sequence.
SOURCE	Tetraodon nigroviridis
ORGANISM	Tetraodon nigroviridis
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Acinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodontidae; Tetraodon.
1	
REFERENCE	
AUTHORS	Roset Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C., Bernot,A., Fizames,C., Wincker,P., Brothier,P., Quetier,F., Saurin,W. and Weissenbach,J.
TITLE	Estimate of human gene number provided by genome-wide analysis using Tetraodon nigroviridis DNA sequence
JOURNAL	Nat. Genet. 25 (2), 235-238 (2000)
MEDLINE	20296833
PUBMED	10835645
2	
REFERENCE	
AUTHORS	Roset Crollius,H., Jaillon,O., Dasilva,C., Ozouf-Costaz,C., Fizames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and Weissenbach,J.
TITLE	Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis
JOURNAL	Genome Res. 10 (7), 939-949 (2000)
MEDLINE	20359837
PUBMED	10899143
3	
REFERENCE	
AUTHORS	Genoscope.
TITLE	Direct Submission
JOURNAL	Submitted (12-APR-2000) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT	This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/Tetraodon.
FEATURES	Location/Qualifiers
source	1..261
	/organism="Tetraodon nigroviridis"
	/mol_type="genomic DNA"
	/db_xref="taxon:99883"
	/clone="232A01"
	/clone_lib="G"
	/note="Genoscope sequence ID : C0AG232AA01SP1-end : PUC-Ori"
ORIGIN	
Query Match	9.8%; Score 39.8; DB 29; Length 261;
Best Local Similarity	51.8%; Pred. No. 35;
Matches	73; Conservative 4; Mismatches 64; Indels 0; Gaps 0;
Qy	39 AGAGAAATCAAAGAGCCCAAGAAAGAAACCCAGATGAGGAAAGAAACGCCAGCAAGAAAT 98
Db	12 AAATGGBNNWGAAAAA 71
Qy	99 GACTACATTTCAAGAGAAACTTCAAGATCAAGATAAGAAAGCCAGAGAGTTTCATCCAC 158
Db	72 AAATAAATGAAAGAAAAAATTAATAAATAAAAAAAAAAATGTAATAAAAAAAAAAAANW 131
Qy	159 TTCTAATCAGGAAACGAGAA 179
Db	132 ATTTAAANCAGNAAAGAAAA 152
RESULT 15	
AU039357	
LOCUS	AU039357 Dictyostelium discoideum SL (H.Uruehahara) Dictyostelium
DEFINITION	dictyostelium cDNA clone SLH629, mRNA sequence.
ACCESSION	AU039357
VERSION	AU039357.1
KEYWORDS	EST.
SOURCE	Dictyostelium discoideum

ORGANISM Dictyostelium discoideum
 Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
 REFERENCE 1 (bases 1 to 372)
 AUTHORS Morio,T., Urushihara,H., Saito,T., Ugawa,Y., Mizuno,H., Yoshida,M.,
 Yoshino,R., Mitra,B.N., Pi,M., Sato,T., Takemoto,K., Yasukawa,H.,
 Williams,J., Maeda,M., Takeuchi,I., Ochiai,H. and Tanaka,Y.
 TITLE The Dictyostelium developmental cDNA project: generation and
 analysis of expressed sequence tags from the first-finger stage of
 development
 JOURNAL DNA Res. 5 (6), 335-340 (1998)
 MEDLINE 99156227
 PUBMED 10048482
 COMMENT Contact: Hideko Urushihara
 Institute of Biological Sciences
 University of Tsukuba
 1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8572, Japan
 Tel: 81-298-53-4664
 Fax: 81-298-53-6614
 Email: hideko@biol.tsukuba.ac.jp
 PROJECT = 'Dictyostelium discoideum cDNA project in Japan'.
 FEATURES
 source
 1..372
 Location/Qualifiers
 /organism="Dictyostelium discoideum"
 /mol_type="mRNA"
 /strain="AX4"
 /db_xref="taxon:44689"
 /clone="SLH629"
 /dev_stage="slug"
 /clone_lib="Dictyostelium discoideum SL (H.Urushihara)"
 ORIGIN
 Query Match 9.8%; Score 39.6; DB 9; Length 372;
 Best Local Similarity 54.9%; Pred. NO. 40;
 Matches 78; Conservative 0; Mismatches 64; Indels 0; Gaps 0;
 QY 39 AGAGAAATCAAAGAGAGCCCAAGAGAAACCCAGATCAGGAAGAGAAACCGCAGGAAAT 98
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 140 AGAAAAGGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAG 199
 QY 99 GACTACATTTGAAGAAAACCTTCAGATCAAGATAAGAAAAGCCCAAGAGTTTCATCCAC 158
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 200 AGAAAAGGATAAGAAAAGAAAAGAAAAGAAAAGAAAAGATAAGAAAAGATAAGAAA 259
 QY 159 TTCTAATCAGGAAAACGAGAAAT 180
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 260 AGATAATTAAGAAAAGATAAT 281

Search completed: April 23, 2004, 02:33:55
 Job time : 2027 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 22, 2004, 17:20:54 ; Search time 291 Seconds
(without alignments)
5912.441 Million cell updates/sec

Title: US-10-005-907-1_COPY_25_429
Perfect score: 405
Sequence: 1 atgggaattatctctctgog.....atgaagttgtgtttccacac 405

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 4428564

Minimum DB seq length: 0
Maximum DB seq length: 405

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38.8	9.6	172	5 AAF24365	Aaf24365 Retrovira
2	38.6	9.5	403	7 ABT22783	Abt22783 Breast ca
3	37.2	9.2	283	2 AA212627	Aa212627 Human gen
4	37.2	9.2	384	5 ABV18980	Abv18980 Human pro
5	36.6	9.0	390	4 AAI86998	Aai86998 Human pol
6	36.4	9.0	325	5 AAS29049	Aas29049 CDNA enco
7	36.4	9.0	325	6 ABS68189	Abs68189 CDNA enco
8	36.4	9.0	325	9 ADC25183	Adc25183 Human CDN
9	36.2	8.9	346	5 ABV54289	Abv54289 Human pro
10	36.2	8.9	396	4 AAI85197	Aai85197 Human pol
11	36	8.9	343	7 ABX39335	Abx39335 Bovine ES
12	36	8.9	381	6 AB178974	Ab178974 Human ova
13	36	8.9	382	5 ABV54300	Abv54300 Human pro
14	35.8	8.8	317	5 ABV58181	Abv58181 Human pro
15	35.6	8.8	330	5 ABV18573	Abv18573 Human pro
16	35.4	8.7	329	5 ABV50079	Abv50079 Human pro
17	35.4	8.7	332	7 ABX46789	Abx46789 Bovine ES
18	35.4	8.7	341	7 ABX47137	Abx47137 Bovine ES
19	35.2	8.7	285	4 AAI80160	Aai80160 Human pol
20	35	8.6	209	5 ABV57456	Abv57456 Human pro
21	35	8.6	233	5 ABV58773	Abv58773 Human pro
22	34.8	8.6	338	5 ABV48535	Abv48535 Human pro
23	34.8	8.6	385	5 ABV54241	Abv54241 Human pro

c	24	34.8	8.6	402	4	AAI82045	Aai82045 Human pol
	25	34.4	8.5	195	7	ACA25512	Aca25512 Prokaryot
	26	34.2	8.4	252	5	ABV58215	Abv58215 Human pro
	27	34.2	8.4	290	6	ABL87783	Ab187783 Human ova
	28	34.2	8.4	299	7	ABX38127	Abx38127 Bovine ES
	29	34.2	8.4	309	7	ABX41877	Abx41877 Bovine ES
c	30	34.2	8.4	393	4	AAI11189	Aai11189 Human bre
	31	34.2	8.4	394	4	AAI87163	Aai87163 Human pol
	32	34.2	8.4	403	4	AAI87975	Aai87975 Human pol
	33	34.2	8.4	403	4	AAI83197	Aai83197 Human pol
	34	34	8.4	318	8	ADA58930	Ada58930 Maize suc
	35	34	8.4	366	4	AAI83212	Aai83212 Human pol
	36	34	8.4	392	7	ABX39643	Abx39643 Bovine ES
	37	33.8	8.3	208	3	ACQ98557	Aac98557 Human col
	38	33.8	8.3	224	5	ABV57906	Abv57906 Human pro
	39	33.8	8.3	262	5	ABV58178	Abv58178 Human pro
c	40	33.8	8.3	332	7	ABX39719	Abx39719 Bovine ES
	41	33.8	8.3	347	5	ABV58528	Abv58528 Human pro
	42	33.8	8.3	355	4	AAI81491	Aai81491 Human pol
c	43	33.8	8.3	357	4	AAK90758	Aak90758 Human dig
	44	33.8	8.3	357	4	AAK90757	Aak90757 Human dig
c	45	33.8	8.3	357	4	AAK90760	Aak90760 Human dig

ALIGNMENTS

RESULT 1

AAF24365
ID AAF24365 standard; DNA; 172 BP.
XX
AC AAF24365;
XX
DT 24-APR-2001 (first entry)
XX
DE Retroviral recombination assay coding sequence fragment #10.
XX
KW Retroviral recombination assay; gene therapy vector; viral vector; gag;
KW pol; replication; HIV; ds.
XX
OS Unidentified.
XX
PN WO200104360-A2.
XX
PD 18-JAN-2001.
XX
PF 06-JUL-2000; 2000WO-US018597.
XX
PR 09-JUL-1999; 99US-0143015P.
PR 10-NOV-1999; 99US-0164626P.
XX
XX (UABR-) UAB RES FOUND.
XX (TRAN-) TRANZYME INC.
XX Kappes JC, Wu X, Wakefield J;
XX WFI; 2001-091927/10.
XX
XX Retroviral recombination assays, systems and cells, useful for evaluating the risk of producing a replication-competent retrovirus from a retroviral-based vector.
XX
XX Example 5; Fig 14; 88pp; English.
XX
XX The present invention describes a method for detecting a retroviral genetic recombinant having gag and pol functions, involving a cell suspected of having a recombinant and propagating the recombinant in the presence of any necessary helper functions. This is useful as it enables the detection of vectors which are able to replicate in the host cell, and allows the production of vectors suitable for gene therapy. The present sequence is a DNA fragment used to demonstrate the method of the invention

CC cancerous state of a mammalian cell, comprising detecting at least one
CC differentially expressed gene product in a test sample from a cell
CC suspected of being cancerous, where the gene product is encoded by one of
CC the 5248 polynucleotide sequences given in AA212532 to AA217779. The
CC polynucleotides can be used as a source of primers and probes, which can
CC be used for a variety of purpose, e.g. detection of expression levels,
CC mapping, tissue typing or profiling, forensic, genetic analysis and
CC detection of polymorphisms. Polypeptides encoded by the polynucleotides
CC can be used for raising antibodies for experimental, diagnostic and
CC therapeutic purposes. The polynucleotides may also be used to construct
CC arrays for diagnostics (which may be used to determine function of an
CC encoded protein); and to detect differences in expression levels between
CC two cells (e.g. to identify abnormal or diseased tissue in a human, to
CC identify a genetic predisposition or susceptibility to a disease such as
CC cancer). The polynucleotides of the invention are especially used in the
CC diagnosis, prognosis and management of colorectal cancer, breast cancer,
CC and lung cancer. The polynucleotides can also be used to screen for
CC peptide analogues and antagonists
XX
SQ Sequence 283 BP; 20 A; 27 C; 34 G; 126 T; 0 U; 76 Other;

Query Match 9.2%; Score 37.2; DB 2; Length 283;
Best Local Similarity 42.9%; Pred. No. 1;
Matches 66; Conservative 0; Mismatches 88; Indels 0; Gaps 0;
QY 41 AGAATCAAGAGAGCCCAAGAAAGGAAACCCAGATGAGGAAAGAAACCGCAGGAAATGA 100
Db 255 AAANGNAGNANCAA 196
QY 101 CTACATTTGAAGAAACTTCAGATCAAGTACAGTAAAGCCAGAGTTTCACACTT 160
Db 195 AAATAAANGNAAAT 136
QY 161 CTAATCAGGAAAGCAGAAATGGCAGTGGTCTGA 194
Db 135 AAANNCAGNTANNGTGA 102

RESULT 4
ABV18980
ID ABV18980 standard; cDNA; 384 BP.
XX
AC ABV18980;
XX
DT 13-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 18971.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
FN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US005171.
XX
PR 17-FEB-2000; 2000US-0183319P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JE;
XX
DR WPI; 2001-662795/76.
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of

PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
PS Claim 1; Page 3116; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 384 BP; 163 A; 57 C; 23 G; 82 T; 0 U; 59 Other;

Query Match 9.2%; Score 37.2; DB 5; Length 384;
Best Local Similarity 50.6%; Pred. No. 1.1;
Matches 78; Conservative 0; Mismatches 76; Indels 0; Gaps 0;
QY 8 ATTATCTCTCGGAAACTCAGTTCCTCGGAGAGATCAAAAGAGCCCAAGAAAGAA 67
Db 1 AGTACTTCTAGAAATTAATTAAGCGGGGGGAGACAAAAAAGAAAAAAGAAAAA 60
QY 68 ACCCAGATCAGGAAAGAAACCGCAGGAAATGACTACATTTGAAAGAAAACTTCAAGATC 127
Db 61 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 120
QY 128 AAGATAAGAAAGCCCAAGAGTTTCATCCACTTC 161
Db 121 AANANAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 154

RESULT 5
AA186998
ID AA186998 standard; cDNA; 390 BP.
XX
AC AA186998;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 7058.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation; ss.
XX
OS Homo sapiens.
XX
FN WO200164835-A2.
XX
PD 07-SEP-2001.
XX
PF 26-FEB-2001; 2001WO-US004927.
XX
PR 28-FEB-2000; 2000US-00515126.
PR 18-MAY-2000; 2000US-00577409.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-514838/56.
DR P-PSDB; AAO07067.
XX
PT Isolated nucleic acids and polypeptides, useful for preventing diagnosing
PT and treating e.g. leukemia, inflammation and immune disorders.
XX

PS Claim 1; SEQ ID NO 7058; 1399pp + Sequence Listing; English.

XX The invention relates to human polynucleotides (AAI79941-AAI93841) and

CC the encoded proteins (AAO0010-AAO1310) that exhibit activity relating to

CC cytokine, cell proliferation or cell differentiation or which may induce

CC production of other cytokines in other cell populations. The

CC polynucleotides and polypeptides are useful in gene therapy, vaccines or

CC peptide therapy. The polypeptides have various cytokine-like activities,

CC e.g. stem cell growth factor activity, haematopoiesis regulating

CC activity, tissue growth factor activity, immunomodulatory activity and

CC activin/inhibin activity and may be useful in the diagnosis and/or

CC treatment of cancer, leukaemia, nervous system disorders, arthritis and

CC inflammation. Note: The sequence data for this patent did not form part

CC of the printed specification, but was obtained in electronic format

CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 390 BP; 229 A; 48 C; 49 G; 57 T; 0 U; 7 Other;

SEQ

Query Match 9.0%; Score 36.6; DB 4; Length 390;

Best Local Similarity 54.0%; Pred. No. 1.7;

Matches 75; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 39 AGAGATCAAAAGAGCCCAAGAGAAAGCCAGATGAGGAAAGAAACGGCAGGAAT 98

DG 135 AA 194

QY 99 GACTACATTGCAAGAAACTTCAAGATCAAGATAAGAAAGCCAGAGTTTCATCCAC 158

DG 195 AA 254

QY 159 TTCTAATCAGGAACGAG 177

DG 255 AATAAATAAAAAAAAAAGGGG 273

RESULT 6

ID AAS29049

XX AAS29049 standard; cDNA; 325 BP.

AC AAS29049;

XX

DT 21-NOV-2001 (first entry)

XX

DE cDNA encoding for human DNA-binding protein #20.

DE

DE Human; DNA-binding protein; histone; chromo domain protein;

KW Chromatin organisation modifier; Y-box binding protein; DNA organisation;

KW gene transcription; malignant disease; autoimmune disorder;

KW rheumatic disease; genetic abnormality; infectious disease;

KW neurological disorder; gene therapy; immunomodulatory; anti-HIV;

KW anti rheumatic; anti microbial; cytostatic; ss.

XX

OS Homo sapiens.

XX

FN WO20015162-A1.

XX

PD 02-AUG-2001.

XX

PF 17-JAN-2001; 2001WO-US001305.

XX

XX 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216880P.

PR

PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.

PR 26-JUL-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225213P.

PR 14-AUG-2000; 2000US-0225214P.

PR 14-AUG-2000; 2000US-0225266P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.

PR 14-AUG-2000; 2000US-0225757P.

PR 14-AUG-2000; 2000US-0225758P.

PR 14-AUG-2000; 2000US-0225759P.

PR 18-AUG-2000; 2000US-0226279P.

PR 22-AUG-2000; 2000US-0226681P.

PR 22-AUG-2000; 2000US-0226689P.

PR 23-AUG-2000; 2000US-0227182P.

PR 23-AUG-2000; 2000US-0227009P.

PR 30-AUG-2000; 2000US-0228924P.

PR 01-SEP-2000; 2000US-0229287P.

PR 01-SEP-2000; 2000US-0229343P.

PR 01-SEP-2000; 2000US-0229344P.

PR 01-SEP-2000; 2000US-0229345P.

PR 05-SEP-2000; 2000US-0229509P.

PR 05-SEP-2000; 2000US-0229513P.

PR 06-SEP-2000; 2000US-0230437P.

PR 06-SEP-2000; 2000US-0230438P.

PR 08-SEP-2000; 2000US-0231242P.

PR 08-SEP-2000; 2000US-0231243P.

PR 08-SEP-2000; 2000US-0231244P.

PR 08-SEP-2000; 2000US-0231413P.

PR 08-SEP-2000; 2000US-0231414P.

PR 08-SEP-2000; 2000US-0232080P.

PR 08-SEP-2000; 2000US-0232081P.

PR 12-SEP-2000; 2000US-0231968P.

PR 14-SEP-2000; 2000US-0232397P.

PR 14-SEP-2000; 2000US-0232398P.

PR 14-SEP-2000; 2000US-0232399P.

PR 14-SEP-2000; 2000US-0232400P.

PR 14-SEP-2000; 2000US-0232401P.

PR 14-SEP-2000; 2000US-0232403P.

PR 14-SEP-2000; 2000US-0233063P.

PR 14-SEP-2000; 2000US-0233064P.

PR 14-SEP-2000; 2000US-0233065P.

PR 21-SEP-2000; 2000US-0234223P.

PR 21-SEP-2000; 2000US-0234274P.

PR 25-SEP-2000; 2000US-0234997P.

PR 25-SEP-2000; 2000US-0234998P.

PR 26-SEP-2000; 2000US-0235484P.

PR 27-SEP-2000; 2000US-0235834P.

PR 27-SEP-2000; 2000US-0235836P.

PR 29-SEP-2000; 2000US-0236327P.

PR 29-SEP-2000; 2000US-0236367P.

PR 29-SEP-2000; 2000US-0236368P.

PR 29-SEP-2000; 2000US-0236369P.

PR 29-SEP-2000; 2000US-0236370P.

PR 02-OCT-2000; 2000US-0236802P.

PR 02-OCT-2000; 2000US-0237037P.

PR 02-OCT-2000; 2000US-0237038P.

PR 02-OCT-2000; 2000US-0237039P.

PR 02-OCT-2000; 2000US-0237040P.

PR 13-OCT-2000; 2000US-0239355P.

PR 13-OCT-2000; 2000US-0239357P.

PR 20-OCT-2000; 2000US-0240960P.

PR 20-OCT-2000; 2000US-0241221P.

PR 20-OCT-2000; 2000US-0241785P.

PR 20-OCT-2000; 2000US-0241786P.

PR 20-OCT-2000; 2000US-0241787P.

PR 20-OCT-2000; 2000US-0241808P.

PR 20-OCT-2000; 2000US-0241809P.

20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 01-DEC-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-02559678P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-465557/50.
XX P-PSDB; AAU18173.
XX
XX Nucleic acid molecules encoding human secreted chromosomal binding
XX proteins, used in preventing, treating or ameliorating a disorder, e.g.
XX Alzheimer's and Parkinson's diseases and cancers.
XX
XX Claim 4; SEQ ID NO 30; 561bp; English.
XX
XX The present invention relates to the isolation of novel DNA-binding
XX proteins (AAU18154-AAU18281), and cDNA and genomic sequences encoding for
XX these proteins. DNA-binding proteins such as histones, chromatin
XX organisation modifier domain proteins, and Y-box binding proteins may
XX contribute to diseases resulting from aberrant DNA organisation and/or
XX gene transcription. The sequences of the invention are useful in
XX screening assays to identify antagonists and/or agonists that may enhance
XX or block activities mediated by DNA-binding proteins. Blockers of DNA-
XX binding proteins may be useful in treating disorders such as malignant

CC diseases (e.g. cancer), autoimmune disorders (e.g. diabetes mellitus),
CC rheumatic diseases (e.g. rheumatoid arthritis), genetic abnormalities
CC (e.g. cystic fibrosis), infectious diseases (e.g. HIV) and neurological
CC disorders (e.g. Alzheimer's disease). The polynucleotide sequences of the
CC invention may also be used in gene therapy. AAS29030-AAS29157 represent
CC cDNA sequences encoding for novel DNA-binding proteins. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 325 BP; 174 A; 38 C; 36 G; 77 T; 0 U; 0 Other;
SQ
Query Match 9.0%; Score 36.4; DB 5; Length 325;
Best Local Similarity 53.5%; Pred. No. 1.8;
Matches 76; Conservative 0; Mismatches 66; Indels 0; Gaps 0;
QY 6 AAATTATCTCTCGGAAACTCAGTTGCTGGGAGAGAAATCAAAAGAGCCCAAGAGG 65
DB 182 AGATTATCTATTGCAAAAAGATATTTCACACCTAAATAAAAAAAAAAAAAA 241
QY 66 AAACCCAGATGAGGAAGAAACGGCAGGAAATGACTACATTGTAAGAAACTTCAAGA 125
DB 242 AA 301
QY 126 TCAAGATTAAGAAAGCCAAAGAA 147
DB 302 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 323
RESULT 7
ABS68189
ID ABS68189 standard; cDNA; 325 BP.
XX
XX AC ABS68189;
XX
XX DT 18-NOV-2002 (first entry)
XX
XX DE cDNA encoding human DNA-binding protein #20.
XX
XX KW Human; DNA-binding protein; B cell immunodeficiency; autoimmune disorder;
XX severe combined immunodeficiency; rheumatoid arthritis; Crohn's disease;
XX diabetes mellitus; allergy; asthma; inflammatory condition; thrombosis;
XX graft-versus-host disease; blood-related disorder; atherosclerosis;
XX hyperproliferative disorder; cancer; renal disorder; arrhythmia;
XX acute glomerulonephritis; cardiovascular disorder; respiratory disorder;
XX Goodpasture's syndrome; neurological disorder; Alzheimer's disease;
XX Parkinson's disease; endocrine disorder; Addison's disease; gene;
XX reproductive system disorder; endometriosis; infectious disease;
XX viral infection; bacterial infection; fungal infection; vaccine;
XX gastrointestinal disorder; multiple sclerosis; gene therapy; ss.
XX
XX OS Homo sapiens.
XX
XX PN US2002102638-A1.
XX
XX PD 01-AUG-2002.
XX
XX PF 17-JAN-2001; 2001US-00764846.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
XX 28-JUN-2000; 2000US-0214886P.
XX 07-JUL-2000; 2000US-0216847P.
XX 07-JUL-2000; 2000US-0216880P.
XX 11-JUL-2000; 2000US-0217487P.
XX 11-JUL-2000; 2000US-0217496P.
XX 14-JUL-2000; 2000US-0218290P.
XX 26-JUL-2000; 2000US-0220963P.
XX 26-JUL-2000; 2000US-0220964P.
XX 14-AUG-2000; 2000US-0224518P.
XX 14-AUG-2000; 2000US-0224519P.
XX 14-AUG-2000; 2000US-0225267P.
XX 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 22-AUG-2000; 2000US-0226868P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 08-SEP-2000; 2000US-0231413P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 17-NOV-2000; 2000US-0249299P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
XX

(ROSE/) ROSEN C A.

(RUBE/) RUBEN S M.

(BARA/) BARASH S C.

Rosen CA, Ruben SM, Barash SC;

WPI; 2002-690611/74.

P-PSDB; ABC92594.

Novel DNA-binding protein useful for diagnosis, prognosis, prevention and treatment of immune, hyperproliferative, respiratory, cardiovascular, reproductive, endocrine, gastrointestinal and neurological disorders.

Claim 1; SEQ ID NO 30; 225pp; English.

The present invention relates to a new DNA-binding protein. The invention is useful in treating, preventing, diagnosing and/or prognosing immunodeficiencies (e.g. B cell immunodeficiencies, severe combined immunodeficiencies), autoimmune disorders (rheumatoid arthritis, multiple sclerosis, diabetes mellitus), allergic reactions and conditions (e.g. asthma), inflammatory conditions, graft-versus-host disease, blood-related disorders (thrombosis, atherosclerosis), hyperproliferative disorders (e.g. cancer), renal disorders (e.g. acute glomerulonephritis), cardiovascular disorders (e.g. arrhythmia), respiratory disorders (Goodpasture's syndrome), neurological disorders (e.g. Alzheimer's disease, Parkinson's disease), endocrine disorders (e.g. Addison's disease), reproductive system disorders (e.g. endometriosis), infectious diseases (e.g. viral, bacterial or fungal infections) and gastrointestinal disorders (e.g. Crohn's disease). The invention is also useful to stimulate neuronal growth and treat, prevent, and/or diagnose neuronal damage which occurs in certain neuronal disorders or neurodegenerative conditions. The present nucleic acid sequence encodes a human DNA-binding protein of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from USPTO at <http://seqdata.uspto.gov/sequence>

XX SQ Sequence 325 BP; 174 A; 38 C; 36 G; 77 T; 0 U; 0 Other;
Query Match 9.0%; Score 36.4; DB 6; Length 325;
Best Local Similarity 53.5%; Pred. No. 1.8;
Matches 76; Conservative 0; Mismatches 66; Indels 0; Gaps 0;
QY 6 AAATTATCTCTCGGAAACTCAGTTGCTGGGAGAGATCAAAAGAGCCCAAGAGG 65
DB 182 AGATTATCTTATTCGAAAGAGATATTTCAAACTCAAAAAA 241
QY 66 AAACCCAGATGAGGAGAAACGCGAGGAATGACTACTTGAAGAGAACTTCAAGA 125
DB 242 AA 301
QY 126 TCAGATTAAGAAAGCCAGAA 147
DB 302 AAAAAAAAAAAAAAAAAAAAAAAAAA 323
RESULT 8
ADC25183
ID ADC25183 standard; cDNA; 325 BP.
XX
AC ADC25183;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human cDNA from extracellular matrix gene 20.
XX
KW Extracellular matrix protein; cytostatic; antibacterial; virucide;
KW neuroprotective; gynaecological; gastrointestinal-Gen; cardiant;
KW cardiovascular-Gen; nephrotropic; antiinflammatory; muscular-Gen;
KW respiratory-Gen; immunosuppressive; cerebroprotective; vasotropic;
KW nootropic; antiallergic; cancer; bacterial infection; viral infection;
KW neural disorder; immune system disorder; blood disorder;
KW muscular disorder; reproductive disorder; gastrointestinal disorder;
KW pulmonary disorder; cardiovascular disorder; renal disorder;
KW inflammatory disorder; proliferative disorder; human; gene therapy; ss;
XX
XX Homo sapiens.
XX
XX US2003049650-A1.
XX
XX 13-MAR-2003.
XX
XX 07-MAR-2002; 2002US-00091483.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
XX 24-FEB-2000; 2000US-0184664P.
XX 02-MAR-2000; 2000US-0186350P.
XX 16-MAR-2000; 2000US-0189874P.
XX 17-MAR-2000; 2000US-0190076P.
XX 18-APR-2000; 2000US-0198123P.
XX 19-MAY-2000; 2000US-0205515P.
XX 07-JUN-2000; 2000US-0209467P.
XX 28-JUN-2000; 2000US-0214886P.
XX 30-JUN-2000; 2000US-0215135P.
XX 07-JUL-2000; 2000US-0216647P.
XX 07-JUL-2000; 2000US-0216880P.
XX 11-JUL-2000; 2000US-0217487P.
XX 11-JUL-2000; 2000US-0217496P.
XX 14-JUL-2000; 2000US-0218290P.
XX 26-JUL-2000; 2000US-0220963P.
XX 26-JUL-2000; 2000US-0220964P.
XX 14-AUG-2000; 2000US-0224518P.
XX 14-AUG-2000; 2000US-0224519P.
XX 14-AUG-2000; 2000US-0225213P.
XX 14-AUG-2000; 2000US-0225214P.
XX 14-AUG-2000; 2000US-0225266P.
XX 14-AUG-2000; 2000US-0225267P.

CC assay and identifying the protein in the supernatant having the
CC activity). The nucleic acids and proteins display the following
CC activities Cytostatic, antibacterial, Virucide, Neuroprotective,

Query Match 9.0%; Score 36.4; DB 9; Length 325;
Best Local Similarity 53.5%; Pred. NO. 1.8;
Matches 76; Conservative 0; Mismatches 66; Indels 0; Gaps 0;
QY 6 AATATATCTCTGCGAAACTCAGTTGCTGGGAGAGATCAAAAGAGCCCAAGAGG 65
Db 182 AGATTATCTATTGCAAAAGATATTTCAAACTTAAAAAAGAAAAAAGAAAAA 241
QY 66 AAACCCAGATCAGGAGAAAGAAACGCGAGAAATGCTACATTTCAAAAGAAAACTTCAAGA 125
Db 242 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 301
QY 126 TCAAGATAAGAAAGCCCAAGAA 147
Db 302 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 323

RESULT 9
ABV54289
ID ABV54289 standard; cDNA; 346 BP.
XX
AC ABV54289;
XX
DT 17-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 54280.

XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
XX Homo sapiens.

OS Homo sapiens.
XX
XX WC200160860-A2.
XX
XX 23-AUG-2001.

XX 20-FEB-2001; 2001WO-US005171.
XX
XX 17-FEB-2000; 2000US-0183319P.
XX
XX 16-MAR-2000; 2000US-0189862P.
XX
XX 25-MAY-2000; 2000US-0207454P.
XX
XX 09-JUN-2000; 2000US-0211314P.
XX
XX 18-JUL-2000; 2000US-0213007P.
XX
XX 13-DEC-2000; 2000US-0255281P.

(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Endege WO, Monahan JE;
XX
XX WPI; 2001-662795/76.
XX

Novel isolated nucleic acid molecule associated with cancerous state of
prostate cells and correlating with presence of prostate cancer, useful
for detecting presence of prostate cancer, stage of prostate cancer.

Claim 1; Page 10498; 11750pp; English.

The invention relates to an isolated nucleic acid molecule (I) comprising
a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
specification or its complement. (I) is useful for: (a) assessing whether
a patient is afflicted with prostate cancer; (b) monitoring the
progression of prostate cancer in a patient; (c) assessing the efficacy
of a test compound to inhibit prostate cancer in a patient; (d) assessing
the efficacy of a therapy for inhibiting prostate cancer in a patient;
(e) selecting a composition for inhibiting prostate cancer in a patient;
(f) assessing the prostate cell carcinogenic potential of a compound; (g)
determining whether prostate cancer has metastasized in a patient; (h)
assessing the aggressiveness or indolence of prostate cancer in a patient
XX ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker

XX Sequence 346 BP; 138 A; 85 C; 52 G; 71 T; 0 U; 0 Other;
SQ
Query Match 8.9%; Score 36.2; DB 5; Length 346;
Best Local Similarity 53.1%; Pred. NO. 2.1;
Matches 77; Conservative 0; Mismatches 68; Indels 0; Gaps 0;
QY 14 TCCTCGAAAACTCAGTTGCTGGGAGAGATCAAAAGAGCCCAAGAGAAACCCAG 73
Db 14 TACTTCTAGATTAATTAACGCGGGGAAAAAAGAAAAAAGAAAAAAGAAAAA 73
QY 74 ATGAGGAAAGAAAAACGCGAGGAATGACTACATTTGAAAGAAACTTCAAGATCAAGATA 133
Db 74 AGTAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 133
QY 134 AGAAAGCCCAAGAGTTTCATCCAC 158
Db 134 AAAAAAAGAAAAAAGGGGCCCC 158

RESULT 10
AAI85197
ID AAI85197 standard; cDNA; 396 BP.
XX
AC AAI85197;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 5257.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation; ss.

OS Homo sapiens.
XX
XX WC200164835-A2.
XX
XX 07-SEP-2001.

XX 26-FEB-2001; 2001WO-US004927.
XX
XX 28-FEB-2000; 2000US-00515126.
XX
XX 18-MAY-2000; 2000US-00577409.
XX
XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;
XX
XX WPI; 2001-514838/56.
XX
XX P-PSDB; AAO05266.

Isolated nucleic acids and polypeptides, useful for preventing diagnosing
and treating e.g. leukemia, inflammation and immune disorders.

Claim 1; SEQ ID NO 5257; 1399pp + Sequence Listing; English.

The invention relates to human polynucleotides (AAI79941-AAI93841) and
the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to
cytokine, cell proliferation or cell differentiation or which may induce
production of other cytokines in other cell populations. The
polynucleotides and polypeptides are useful in gene therapy, vaccines or
peptide therapy. The polypeptides have various cytokine-like activities,
e.g. stem cell growth factor activity, haematopoiesis regulating
activity, tissue growth factor activity, immunomodulatory activity and
activin/inhibin activity and may be useful in the diagnosis and/or
treatment of cancer, leukaemia, nervous system disorders, arthritis and
inflammation. Note: the sequence data for this patent did not form part
of the printed specification, but was obtained in electronic format
directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 396 BP; 178 A; 47 C; 100 G; 71 T; 0 U; 0 Other;

The invention relates to a purified nucleic acid molecule associated with lactation or muscle and fat deposition (designated LMFD), derived from cattle, and the LMFD nucleic acid can specifically hybridise to a second nucleic acid molecule comprising any of 15112 nucleotide sequences, appearing as AX334836-ABX49947, or complements of them. Also included are:

- (1) a transformed cell having a nucleic acid comprising an LMFD nucleic acid linked to a promoter and a 3' non-translated sequence that functions in the cell to cause termination of transcription and addition of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
- (2) determining a level or pattern of a molecule in a bovine cell or tissue comprising: (a) incubating a marker nucleic acid (comprising any of the 15112 nucleic acid sequences or its complement or fragment) with a complementary nucleic acid molecule obtained from the bovine cell or tissue, where hybridisation between the marker nucleic acid and the complementary nucleic acid permits the detection of the molecule; and (b)

XX Composition for therapy and diagnosis of ovarian cancer comprising
PT polypeptide of a ovarian tumor polypeptide, polynucleotide encoding
PT polypeptide, antibody specific to polypeptide or T cell expressing
PT polypeptide.
XX
XX Claim 1; SEQ ID NO 1952; 489pp; English.
XX
XX The present invention describes a composition (I) comprising: carriers
CC and immunostimulants; and a polypeptide (II) of a ovarian tumour
CC polypeptide encoded by a polynucleotide (III) having a cDNA sequence (S1)
CC from the 19312 nucleotide sequences as given in ABU77023 to ABU7934,
CC (III) encoding (II) having a sequence (S2), a T cell population of (II),
CC or antigen presenting cells that express (II). (I) has cytostatic
CC activity. An oligonucleotide (IV) that hybridises to (S1) can be used for

CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
XX

SQ Sequence 317 BP; 146 A; 44 C; 30 G; 96 T; 0 U; 1 Other;

Query Match 8.8%; Score 35.8; DB 5; Length 317;
Best Local Similarity 51.2%; Pred. No. 2.6;
Matches 82; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

QY 14 TCCTGCGAAACTCAGTTGCTGCGAGAGAAATCAAAAGAGCCCAAGAAAGGAAACCCAG 73

Db 13 TACTTCTAGATTAAATTAACGCGGGGNGAAAAAATAAAAAAAAAAAAAAAAAAAAAA 72

QY 74 ATGAGGAAAGAAACGCGCAGGAATGACTACATTGGAAGAAACITTCAGATCAAGATA 133

Db 73 AA 132

QY 134 AGAAAGCCAGAGTTTCATCCACTTCTAATCAGGAAAA 173

Db 133 AAAAAAAAAAAAAAATAAATAAATAATTTTAAAAA 172

RESULT 15

ABV18573

ID ABV18573 standard; cDNA; 330 BP.

XX

AC ABV18573;

XX

DT 13-SEP-2002 (first entry)

XX

XX Human prostate expression marker cDNA 18564.

DE

KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

KW

XX pharmacogenomic marker; gene; ss.

XX

OS Homo sapiens.

XX

FN WO200160860-A2.

XX

XX 23-AUG-2001.

XX

XX 20-FEB-2001; 2001WO-US005171.

XX

XX 17-FEB-2000; 2000US-0183319P.

XX

XX 16-MAR-2000; 2000US-0189862P.

XX

XX 25-MAY-2000; 2000US-0207454P.

XX

XX 09-JUN-2000; 2000US-0211314P.

XX

XX 18-JUL-2000; 2000US-0219007P.

XX

XX 13-DEC-2000; 2000US-0255281P.

XX

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX

XX Schlegel R, Endege WO, Monahan JE;

XX

XX WPI; 2001-662795/76.

XX

XX Novel isolated nucleic acid molecule associated with cancerous state of

XX

XX prostate cells and correlating with presence of prostate cancer, useful

XX

XX for detecting presence of prostate cancer, stage of prostate cancer.

XX

XX Claim 1; Page 3058; 11750pp; English.

XX

XX The invention relates to an isolated nucleic acid molecule (I) comprising

XX

XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the

XX

XX specification of its complement. (I) is useful for: (a) assessing whether

XX

XX a patient is afflicted with prostate cancer; (b) monitoring the

XX

XX progression of prostate cancer in a patient; (c) assessing the efficacy

XX

XX of a test compound to inhibit prostate cancer in a patient; (d) assessing

XX

CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
XX

SQ Sequence 330 BP; 174 A; 47 C; 28 G; 57 T; 0 U; 24 Other;

Query Match 8.8%; Score 35.6; DB 5; Length 330;

Best Local Similarity 51.0%; Pred. No. 3.1;

Matches 77; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 8 ATTATCTCTCGGAAACTCAGTTGCTGCGAGAGAAATCAAAAGAGCCCAAGAAAGGAA 67

Db 132 ATTATCTCTTTTAAATACAAATGTCNAAAAAATAAAAAAAAAAAAAAAAAAAAAA 191

QY 68 ACCCAGATGAGGAAAGAAACGCGCAGGAAATGACTACATTTTGAAGAAACTTCAAGATC 127

Db 192 AA 251

QY 128 AAGATAGAAAGCCCAAGAGTTTCATCCAC 158

Db 252 AAAAAAAAAAAAAAAAAANNNNGCCCCCCC 282

Search completed: April 23, 2004, 02:00:01

Job time : 295 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 23, 2004, 02:34:00 ; Search time 292 Seconds
(without alignments)
6253.404 Million cell updates/sec

Title: US-10-005-907-1_COPY_25_429

Perfect score: 405
Sequence: 1 atgggaaatattctctcgcg.....atgaagttgtgtttccacac 405

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2907579 seqs, 2254313464 residues

Total number of hits satisfying chosen parameters: 2611332

Minimum DB seq length: 0
Maximum DB seq length: 405

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38.6	9.5	403	16	US-10-125-968-1156
2	38.4	9.5	347	13	US-10-424-599-101964
3	37.4	9.2	345	13	US-10-424-599-95566
4	36.4	9.0	325	9	US-09-764-846-30
5	36.4	9.0	325	15	US-10-091-483-30
6	36.2	8.9	379	13	US-10-424-599-41873
7	36	8.9	343	9	US-09-960-352-4500
8	36	8.9	381	9	US-09-867-701-1952
9	35.6	8.8	193	9	US-09-969-373-418
10	35.6	8.8	282	13	US-10-424-599-98076
11	35.6	8.8	340	13	US-10-424-599-94516
12	35.4	8.7	332	9	US-09-960-352-11954
13	35.4	8.7	341	9	US-09-960-352-12302
14	35.2	8.7	368	13	US-10-424-599-55157

15	35	8.6	327	13	US-10-424-599-89471	Sequence 89471, A
16	35	8.6	342	13	US-10-424-599-21848	Sequence 21848, A
17	35	8.6	348	13	US-10-424-599-39861	Sequence 39861, A
18	34.8	8.6	234	13	US-10-085-783A-49770	Sequence 49770, A
19	34.8	8.6	234	16	US-10-242-535A-49770	Sequence 49770, A
20	34.8	8.6	270	13	US-10-424-599-139195	Sequence 139195, A
21	34.8	8.6	306	13	US-10-424-599-103588	Sequence 103588, A
22	34.8	8.6	309	13	US-10-424-599-52791	Sequence 52791, A
23	34.6	8.5	183	13	US-10-102-524-1621	Sequence 1621, Ap
24	34.6	8.5	344	13	US-10-424-599-97494	Sequence 97494, A
25	34.6	8.5	358	10	US-09-814-353-4733	Sequence 4733, Ap
26	34.6	8.5	358	10	US-09-814-353-11030	Sequence 11030, A
27	34.6	8.5	391	13	US-10-424-599-48934	Sequence 48934, A
28	34.4	8.5	195	13	US-10-282-122A-13382	Sequence 13382, A
29	34.4	8.5	283	13	US-10-424-599-85844	Sequence 85844, A
30	34.2	8.4	290	9	US-09-867-701-10761	Sequence 10761, A
31	34.2	8.4	299	9	US-09-960-352-3292	Sequence 3292, Ap
32	34.2	8.4	309	9	US-09-960-352-7042	Sequence 7042, Ap
33	34.2	8.4	319	13	US-10-085-783A-15028	Sequence 15028, A
34	34.2	8.4	319	16	US-10-242-535A-15028	Sequence 15028, A
35	34.2	8.4	362	10	US-09-814-353-17367	Sequence 17367, A
36	34.2	8.4	388	13	US-10-424-599-131162	Sequence 131162, A
37	34	8.4	279	13	US-10-424-599-83670	Sequence 83670, A
38	34	8.4	318	10	US-09-237-183A-1084	Sequence 1084, Ap
39	34	8.4	392	9	US-09-960-352-4808	Sequence 4808, Ap
40	33.8	8.3	208	9	US-09-925-299-567	Sequence 567, App
41	33.8	8.3	208	10	US-09-925-299-567	Sequence 567, App
42	33.8	8.3	235	13	US-10-085-783A-22183	Sequence 22183, A
43	33.8	8.3	235	16	US-10-242-535A-22183	Sequence 22183, A
44	33.8	8.3	332	9	US-09-960-352-4884	Sequence 4884, Ap
45	33.8	8.3	341	13	US-10-424-599-25361	Sequence 25361, A

ALIGNMENTS

RESULT 1

US-10-125-968-1156
; Sequence 1156, Application US/10125968
; Publication No. US20030215805A1
; GENERAL INFORMATION:
; APPLICANT: Lallie, James
; APPLICANT: Palermo, Adam
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; APPLICANT: Elias, Josh
; APPLICANT: Mertens, Maureen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-032
; CURRENT APPLICATION NUMBER: US/10/125,968
; CURRENT FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US 60/285,163
; PRIOR FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 1417
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1156
; LENGTH: 403
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 50, 57, 85, 100, 118, 154, 219, 224, 226, 247, 261
; OTHER INFORMATION: n = A,T,C or G
US-10-125-968-1156

Query Match 9.5%; Score 38.6; DB 16; Length 403;
Best Local Similarity 53.6%; Pred. No. 0.65;
Matches 74; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 21 AAACCTCAGTTCCTGGGAGAGAAATCAAGAGCCCAAGAGGAAAGCCAGATGAGCA 80

209	AAAACTCCATNTCATNTNAGAAAAAAGAAAAAAGAAAAAATATATGAGCA	268
Db		
81	AAGAAAACGGCAGGAAATGACTACATTTGAAAGAAAACCTTCAGATCAGATGAAGAAAAG	140
Qy		
269	TCATATCAAAAAGGCAGCCCTAAGAATAAATGAAAAGTTTCACAGAAAAAATATAAAATG	328
Db		
141	CCAAGAAGTTTTCATGCCAC	158
Qy		
329	CAAAATATCCCTTAAACAC	346
Db		

```

RESULT 2
US-10-424-599-101964
; Sequence 101964, Application US/10424599
; Publication NO. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovacic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 101964
; LENGTH: 347
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_6308C.1
US-10-424-599-101964

```

	Query Match	9.5%	Score 38.4	DB 13	Length 347
	Best Local Similarity	52.5%	Pred. No. 0.69		
	Matches 84	Conservative 0	Mismatches 76	Indels 0	Gaps 0
Qy	41	AGRATCAAAAGAACGCCCAAGAAAGGAAACCCAGATGKGGAAAGAAACGCGCAGGAATGA	100		
Db	95	AAATAAAAA	154		
Qy	101	CTACATTTTGAAGAAAACTTCAAGATCAAGATAAGAAAAGCCAGAAAGTTTCATCCACTT	160		
Db	155	ATAAATATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATATAAAAAAAAAACCAACAT	214		
Qy	161	CTAATCAGGAAACGAGAAATGCAGTGGTTCTGAAGAAGT	200		
Db	215	AAATATAAAAAAAAAAAAAAAAAAGGGGAATCCTATAAAGAAAT	254		

```

RESULT 3
US-10-424-599-95566
; Sequence 95566, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 39-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 95566
; LENGTH: 345
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_57307C.1
US-10-424-599-95566

```

	Query Match	9.2%	Score 37.4;	DB 13;	Length 345;
	Best Local Similarity	55.9%;	Pred. No. 1.3;		
	Matches 71;	Conservative 0;	Mismatches 56;	Indels 0;	Gaps 0;
Qy	39	ACAGNATCAAAAGAGCCCAAGAAAGGAACCCAGATGAGGAAGAAAACGCCAGGAAT	98		
Db	32	AAAAAATAATAATAAAA	91		
Qy	99	GACTACATTTGAAAGAAAACCTTCAAGATCAAGATAAGAAAGCCAAAGAGTTTCATCCAC	158		
Db	92	AAATATAAATATAAAAAAAAAAATATATATAAATAATAAATAATTATATTCAA	151		
Qy	159	TTCTAAT	165		
Db	152	TTGAATT	158		

```

RESULT 4
US-09-764-846-30
; Sequence 30, Application US/09764846
; Patent No. US20020102638A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT312
; CURRENT APPLICATION NUMBER: US/09/764,846
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
; LENGTH: 325
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-846-30

```

	Query Match	9.0%;	Score 36.4;	DB 9;	Length 325;
	Best Local Similarity	53.5%;	Pred. No. 2.5;		
	Matches	76;	Conservative	0;	Mismatches 66; Indels 0; Gaps 0
Qy	6	AAATTATCTCCTGCGAAAACCTCAGTTCGCTGGGAGAGATCAAAAGAGGCCCAAGAAAGG	65		
Db	182	AGATTATCTATTGCAAAAAGATATTTCAAACCTTAAAAAATAAAAAAAAAAAAAAAAAA	241		
Qy	66	AAACCCAGATGAGGAAAGAAACGCGAGGAATTCATCTACATTTGCAAGAGAAAACCTTCAAGA	125		
Db	242	AAAAAATAA	301		
Qy	126	TCAGATAGAAAAGCCCAAGAA	147		
Db	302	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	323		

```

RESULT 5
US-10-091-483-30
; Sequence 30, Application US/10091483
; Publication No. US20030049650A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PZ12C1
; CURRENT APPLICATION NUMBER: US/10/091,483
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 348
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
; LENGTH: 325
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-091-483-30

```



```
Query Match      9.0%; Score 36.4; DB 15; Length 325;
Best Local Similarity 53.5%; Pred. No. 2.5;
Matches 76; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY      6 AAATTATCTCTCGGAAACTCAGTTGCTGGGAGAGATCAAAAGAGCCCAAGAAAGG 65
Db      182 AGATTATCTATTGCAAAAGATATTTCAAACCTAAAAAAGAAAAAAGAAAAA 241

QY      66 AAACCCAGATCAGAAAGAAACGGCAGGAATGACTTTCAGAGAAACTTCAAGA 125
Db      242 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 301

QY      126 TCAAGATAAGAAAGCCCAAGAA 147
Db      302 AAAAAAAGAAAAAAGAAAAAAGAAAAA 323

RESULT 6
US-10-424-599-41873
; Sequence 41873, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 41873
; LENGTH: 379
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_137813C.1
US-10-424-599-41873

Query Match      8.9%; Score 36.2; DB 13; Length 379;
Best Local Similarity 53.1%; Pred. No. 3.1;
Matches 77; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY      39 AGAGATCTAAGAGAGCCCAAGAAAGGAAACCCAGATGAGGAAGAAAGAAACGGCAGGAAT 98
Db      53 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 112

QY      99 GACTACATTTGAAAGAAACTTCAAGATCAAGATCAAGAAAGCCCAAGAAAGTTTCATCCAC 158
Db      113 AAAAAAAGAAAGAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 172

QY      159 TTCTAATCAGGAAACGAGATGGC 183
Db      173 AAAAAAAGAAAAAAGAAAGGGGGGC 197

RESULT 7
US-09-960-352-4500
; Sequence 4500, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 4500
```

```
; LENGTH: 343
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 20-BOWMS1-017-Q1-E1-E11
US-09-960-352-4500

Query Match      8.9%; Score 36; DB 9; Length 343;
Best Local Similarity 55.6%; Pred. No. 3.4;
Matches 69; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY      36 GGGAGAGATCAAAAGAGCCCAAGAAAGGAAACCCAGATGAGGAAGAAAGAAACGGCAGGA 95
Db      71 GGGATAAAAAAGAAAAAAGAAAAAAGTCTCAAAAAGAGAAAAAAGAAAAAAGAAAAAACA 130

QY      96 AATGACTACATTTGAAGAAAACTTCAAGATCAAGATCAAGAAAGCCCAAGAAAGTTTCATC 155
Db      131 AAAGAAAAAAGAAAAAAGAAAAAAGGAAGTAAAAAATCAAAAAAAGAAAAAAGGGGGCCACC 190

QY      156 CACT 159
Db      191 CAAT 194

RESULT 8
US-09-867-701-1952/C
; Sequence 1952, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 1952
; LENGTH: 381
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(381)
; OTHER INFORMATION: n = A,T,C or G
US-09-867-701-1952

Query Match      8.9%; Score 36; DB 9; Length 381;
Best Local Similarity 62.0%; Pred. No. 3.6;
Matches 57; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY      95 AAATGACTACATTTGAAAGAAAACTTCAAGATCAAGATCAAGAAAGCCCAAGAAAGTTTCAT 154
Db      156 AATTAAGTCTGAGAAATATTTTCAGAAAGAAAGAAATAGACAAAGCTGACTAAA 97

QY      155 CCACTTCTAATCAGGAAACGAGAAATGGCAGT 186
Db      96 CAATATGATGATGAAATAGGTAGGCACT 65

RESULT 9
US-09-969-373-418
; Sequence 418, Application US/09969373
; Patent No. US2002013852A1
; GENERAL INFORMATION:
; APPLICANT: Effertz, Roger J.
; APPLICANT: Hauge, Brian M.
; TITLE OF INVENTION: Soybean SSRs and Methods of Genotyping
; FILE REFERENCE: 38-10(52679)A
; CURRENT APPLICATION NUMBER: US/09/969,373
; CURRENT FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US 09/754,853
```

```
; PRIOR FILING DATE: 2001-01-05
; PRIOR APPLICATION NUMBER: US 09/760,427
; PRIOR FILING DATE: 2001-01-13
; PRIOR APPLICATION NUMBER: US 09/855,768
; PRIOR FILING DATE: 2001-05-15
; NUMBER OF SEQ ID NOS: 4593
; SEQ ID NO 418
; LENGTH: 193
; TYPE: DNA
; ORGANISM: Glycine max
US-09-969-373-418

Query Match      8.8%; Score 35.6; DB 9; Length 193;
Best Local Similarity 57.0%; Pred. No. 3.4;
Matches 65; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 67 AACCCAGATGAGGAAGAAAGCGGAGGAATGACTACATTTGAAAGAAAATCTCAAGAT 126
Db 60 ATCCATTAGAGTAAAGAAACAGAGAAATAATCAACTTAAATTAACATAATCT 119

QY 127 CAAGTAAAGAAAGCCAGAGAGTTTCATCCACTTCTATATCAGGAAACGAGAT 180
Db 120 CAACATATATAACAAATATTATTATTATAATTATAAACAACAAATGCAAT 173

RESULT 10
US-10-424-599-98076
; Sequence 98076, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 98076
; LENGTH: 282
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_59576C.1
US-10-424-599-98076

Query Match      8.8%; Score 35.6; DB 13; Length 282;
Best Local Similarity 54.8%; Pred. No. 4.1;
Matches 71; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 43 AATCAAAAGAGCCCAAGAAAGAAACCCAGATGAGGAAAGAAAGCGGAGGAATGACT 102
Db 71 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 130

QY 103 ACATTTGAAGAAATCTCAAGTCAAGATAGAGAAAGCCAGAGAGTTTCATCCACTTCT 162
Db 131 AGAAATATAACAAATATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAT 190

QY 163 AATCAGAGAA 172
Db 191 AATGAAAGAA 200

RESULT 11
US-10-424-599-94516
; Sequence 94516, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 94516
; LENGTH: 340
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_56359C.1
US-10-424-599-94516

Query Match      8.8%; Score 35.6; DB 13; Length 340;
Best Local Similarity 51.2%; Pred. No. 4.4;
Matches 83; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 41 AGAATCAAAAGAGCCCAAGAAAGAAACCCAGATGAGGAAAGAAAGCGGAGGAATGA 100
Db 97 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 156

QY 101 CTACATTTGAAGAAATCTCAAGTCAAGATCAAGATAGAGAAAGCCAGAGAGTTTCATCCACTT 160
Db 157 TAAAAACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACATACAAAAAATAAATAACAA 216

QY 161 CTAATCAGGAAACGAGAGATGCGAGTGGTCTTGAAGAGTGT 202
Db 217 CACATTATAAATAAATAAAGCGCGGAGTTTAAAAAGTTT 258

RESULT 12
US-09-960-352-11954/c
; Sequence 11954, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Ningbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 11954
; LENGTH: 332
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 51-LIB3058-058-Q1-K1-E4
US-09-960-352-11954

Query Match      8.7%; Score 35.4; DB 9; Length 332;
Best Local Similarity 53.2%; Pred. No. 5;
Matches 75; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 39 AGAGATCAAAAGAGCCCAAGAAAGAAACCCAGATGAGGAAAGAAAGCGGAGGAAT 98
Db 301 AAAAAAGAAAGAAAGATATAAAAAAAAAAAAAAAAAAGAAATAAAGAAAGAAAGAAAT 242

QY 99 GACTACATTTGAAAGAAAGAACTTCAAGATCAAGATAGAGAAAGCCAGAGAGTTTCATCCAC 158
Db 241 AAAAAAAAAAGAAAGATATAAAAAAAAAAAAAAAAAAAAAAAAAAGAGAGAGAGAGAA 182

QY 159 TTCTAATCAGGAAACGAGAA 179
Db 181 AAGAGAAAGAGAAACAAAAA 161

RESULT 13
US-09-960-352-12302/c
```

; Sequence 12302, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Mengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960.352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 12302
; LENGTH: 341
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 53-BOVMS1-009-Q1-E1-F2
US-09-960-352-12302

Query Match 8.7%; Score 35.4; DB 9; Length 341;
Best Local Similarity 53.2%; Pred. No. 5.1;
Matches 75; Conservative 0; Mismatches 66; Indels 0; Gaps 0;
QY 39 ACAGATCAAAAGAGCCCAAGAGAAACCCAGATGAGGAAAGAAACGGCAGGAAAT 98
DB 267 AAAAAAGAAAAACAAAAAGAAAAACAAAAAGAAAAACAAAAAGAAAAACAAAA 208
QY 99 GACTACATTTGAAGAAACCTTCAGATCAAGTAAGAAAGCCCAAGAGTTTCATCCAC 158
DB 207 AAAAAAGAAAAACAAAAAGAAAGACAAAAAGAAAAACAAAAAGAAAAACAAAA 148
QY 159 TTCTAATCAGGAAACGAGAA 179
DB 147 AAAAAAGAAAAACAAAA 127

RESULT 14
US-10-424-599-55157
; Sequence 55157, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 55157
; LENGTH: 368
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(368)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_20818C.1
US-10-424-599-55157

Query Match 8.7%; Score 35.2; DB 13; Length 368;
Best Local Similarity 51.2%; Pred. No. 6;
Matches 82; Conservative 0; Mismatches 78; Indels 0; Gaps 0;
QY 39 AGAGATCAAAAGAGCCCAAGAGAAACCCAGATGAGGAAAGAAACGGCAGGAAAT 98
DB 9 AAAAAAGAAAAACAAAAAGAAAAACAAAAAGAAAAACAAAAAGAAAAACAAAA 68
QY 99 GACTACATTTGAAGAAACCTTCAGATCAAGTAAGAAAGCCCAAGAGTTTCATCCAC 158

DB 59 AAAAAAGAAAAACAAAAAGAAAAACAAAAAGAAAAACAAAAAGAAAAACAAAA 128
QY 159 TTCTAATCAGGAAACGAGAAATGGCAGTGGTTCTGAAGAA 198
DB 129 TTAAAGAAATCAAGTTTACTTCCCGGTCTTCCAAGAA 168

RESULT 15
US-10-424-599-89471
; Sequence 89471, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 89471
; LENGTH: 327
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(327)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_51802C.1
US-10-424-599-89471

Query Match 8.6%; Score 35; DB 13; Length 327;
Best Local Similarity 59.0%; Pred. No. 6.5;
Matches 59; Conservative 0; Mismatches 41; Indels 0; Gaps 0;
QY 18 GCGAAACCTCAGTTGCTGGGAGAGAAATCAAAAGAAAGCCCAAGAGAAACCCAGATCA 77
DB 209 GCGAAACCTCAGTTGCTGGGAGAGAAATCAAAAGAAAGCCCAAGAGAAACCCAGATCA 268
QY 78 GGAAGAAACCGGAGGAAATGACTACATTTGAAAGAAAA 117
DB 269 GGAAGAAACCGGAGGAAATGACTACATTTGAAAGAAAA 308

Search completed: April 23, 2004, 04:03:59
Job time : 298 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 22, 2004, 18:37:20 ; Search time 64 Seconds
(without alignments)
3511.799 Million cell updates/sec

Title: US-10-005-907-1_COPY_25_429
Perfect score: 405
Sequence: 1 atgggaattatctctctg.....atgaattgtgtttccacac 405

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1083176

Minimum DB seq length: 0
Maximum DB seq length: 405

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS.COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	34.8	8.6	399	4	US-09-621-976-8976
2	33.6	8.3	249	4	US-09-621-976-1322
3	33.4	8.2	240	4	US-09-621-976-1324
4	33.3	8.1	400	4	US-08-956-171E-4053
5	31.8	7.9	400	4	US-08-956-171E-3835
6	31.6	7.8	333	3	US-09-018-584A-27
7	31	7.7	240	1	US-08-628-417-6
8	31	7.7	372	3	US-09-018-584A-13
9	30.8	7.6	260	2	US-08-520-678A-29
10	30.8	7.6	260	3	US-08-897-126-29
11	30.8	7.6	377	4	US-08-956-171E-4397
12	30.8	7.6	396	4	US-09-640-173-16
13	30.8	7.6	396	4	US-09-713-550-16
14	30.4	7.5	356	2	US-08-520-678A-22
15	30.4	7.5	356	3	US-08-897-126-22
16	30.2	7.5	144	1	US-08-702-344-26
17	30.2	7.5	147	4	US-09-621-976-10383
18	30	7.4	291	1	US-07-922-723A-7
19	30	7.4	291	1	US-07-799-828C-7
20	30	7.4	291	1	US-08-074-275-7
21	30	7.4	291	1	US-08-480-366-7
22	30	7.4	291	2	US-07-952-277A-7
23	29.8	7.4	251	4	US-08-956-171E-4555
24	29.8	7.4	253	2	US-08-520-678A-25
25	29.8	7.4	253	3	US-08-897-126-25
26	29.6	7.3	248	3	US-09-007-005-32
27	29.6	7.3	248	3	US-09-244-796-32

28	29.6	7.3	277	3	US-09-007-005-3	Sequence 3, Appli
29	29.6	7.3	277	3	US-09-244-796-3	Sequence 3, Appli
30	29.6	7.3	373	2	US-08-557-309B-4	Sequence 4, Appli
31	29.6	7.3	373	3	US-08-834-306-4	Sequence 4, Appli
32	29.6	7.3	373	3	US-08-933-674A-4	Sequence 4, Appli
33	29.6	7.3	373	4	US-09-256-976-4	Sequence 20, Appli
34	29.4	7.3	121	3	US-09-297-535-20	Sequence 4, Appli
35	29.4	7.3	130	4	US-09-621-976-12892	Sequence 12892, A
36	29.4	7.3	147	4	US-09-621-976-10254	Sequence 10254, A
37	29.4	7.3	164	4	US-09-621-976-16692	Sequence 16692, A
38	29.4	7.3	204	3	US-08-910-632-4	Sequence 4, Appli
39	29.4	7.3	204	3	US-08-805-631A-4	Sequence 4, Appli
40	29.4	7.3	204	4	US-09-569-344-4	Sequence 30, Appli
41	29.4	7.3	270	2	US-08-520-678A-30	Sequence 30, Appli
42	29.4	7.3	270	3	US-08-897-126-30	Sequence 20, Appli
43	29.4	7.3	321	3	US-09-018-584A-20	Sequence 18, Appli
44	29.4	7.3	376	2	US-08-623-906A-18	Sequence 5, Appli
45	29.4	7.3	383	3	US-09-155-942-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1

US-09-621-976-8976
; Sequence 8976, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 8976
; LENGTH: 399
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-8976

Query Match	8.6%	Score 34.8;	DB 4;	Length 399;
Best Local Similarity	12.4%	Pred. No. 0.26;		
Matches	30;	Conservative 110;	Mismatches 102;	Indels 0; Gaps 0;
QY	3	GGGAAATTCTCTCGGAAACTCAGTTGCTGGGAGAGATCAAAAGAACCCAGAA	62	
DB	42	KKRRRRRAWWKSKMCKSKSWRSWGMVKMKGRGAASWAGYMSWMTYNTMRMY	101	
QY	63	AGGAAACCCAGATGAGGAGAAAGAACGGCAGAGAAATGACTATTGAAAGAAATCCA	122	
DB	102	YRKYKACTKWRAGNWKAGAWMAYAKWYMAWERTAMKWMKSKSRMRERAWYA	161	
QY	123	AGATCAAGTAAGAAAGCCAGAGTTTCATCCACTTCTAATCAGGAAACAGAGATCG	182	
DB	162	WYTYMARRTWGMGRASCYRGAYMASAGMYWYMYTMRKWMYSAGWSMTRKWTIRCAS	221	
QY	183	CATGGTTCTGAGAGTGTGCTACACTGCTAATATCATCCCCCATCAGAGATCCTC	242	
DB	222	YSCWSSYCMWGAOMWMTKTSRWSYWSYRCKYRSCCCWSMCSYKWTYRSWTCASC	281	
QY	243	CC 244		
DB	282	YY 283		

RESULT 2

US-09-621-976-1322
; Sequence 1322, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:

```

; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 1322
; LENGTH: 249
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 28..249
; US-09-621-976-1322

Query Match      8.3%; Score 33.6; DB 4; Length 249;
Best Local Similarity 56.2%; Pred. No. 0.49; Indels 0; Gaps 0;
Matches 63; Conservative 0; Mismatches 49;

QY 36 GCGAGAGAATCAAAGAGCCCAAGAAAGGAAACCCAGATGAGGAAAGAAACGGCAGGA 95
Db 132 GCGAGAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 191
QY 96 ATGACTACATTGAAAGAAACTTCAAGATCAAGATAAGAAAGCCAGAA 147
Db 192 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 243

RESULT 3
US-09-621-976-1324
; Sequence 1324: Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 1324
; LENGTH: 240
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 28..240
; US-09-621-976-1324

Query Match      8.2%; Score 33.4; DB 4; Length 240;
Best Local Similarity 57.0%; Pred. No. 0.55; Indels 0; Gaps 0;
Matches 61; Conservative 0; Mismatches 45;

QY 36 GCGAGAGAATCAAAGAGCCCAAGAAAGGAAACCCAGATGAGGAAAGAAACGGCAGGA 95
Db 132 GCGAGAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 191
QY 96 ATGACTACATTGAAAGAAACTTCAAGATCAAGATAAGAAAGCC 142
Db 192 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 238

RESULT 4
US-08-956-171E-4053
; Sequence 4053, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch

```

Gil H. Choi
Patrick S. Dillon
Craig A. Rosen
Steven C. Barash
Michael R. Fannon
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5256
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/956,171E
FILING DATE: 20-Oct-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/009,861
FILING DATE: January 5, 1996
APPLICATION NUMBER: 08/781,986
FILING DATE: January 3, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Mark J. Hyman
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB248P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (240) 314-1224
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 4053:
SEQUENCE CHARACTERISTICS:
LENGTH: 400 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 4053:
US-08-956-171E-4053
Query Match 8.1%; Score 33; DB 4; Length 400;
Best Local Similarity 54.5%; Pred. No. 0.93; Mismatches 55; Indels 0; Gaps 0;
Matches 66; Conservative 0;
QY 59 AGAAGAGAAACCCAGATGAGGAGAAAGAAACCGCAGGAAATGACTACATTGTAAGAGAAAC 118
Db 27 ACAAGGATAATTACTTATGCAAGCGGAGGAATCACATGTCTATTACTGAAAAACAAC 86
QY 119 TTCAAGATCAAGATAGAAAACCCAGAGATTTCATCCACTTCTTAATCAGGAAACGAGA 178
Db 87 GTCAGCAACAAGCTGAATTACATAAAAAAATTATGGTCGATGCGGAATGATTTAAGAGGGA 146
QY 179 A 179
Db 147 A 147
RESULT 5
US-08-956-171E-3835
; Sequence 3835, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences

REFERENCE/DOCKET NUMBER: DAM 398-94
TELECOMMUNICATION INFORMATION:
TELEPHONE: 410-671-1158
TELEFAX: 410-671-2534
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 240 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: oligodeoxynucleotide
HYPOTHETICAL: NO
ANTI-SENSE: YES
US-08-628-417-6

Query Match 7.7%; Score 31; DB 1; Length 240;
Best Local Similarity 52.8%; Pred. No. 3.1; Indels 0; Gaps 0;
Matches 67; Conservative 0; Mismatches 60; Indels 0; Gaps 0;
QY 21 AAACTCAGTTGCTGGGAGAGAAATCAAAAGAGCCCAAGAAAGAAACCCAGATGAGGA 80
Db 25 AGAATAATTTTACTTAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 84
QY 81 AAGAAAGCGGAGGAAATGACTTACATTTGAAGAAAGAACTTCAAGATCAAGATAGAAAG 140
Db 85 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 144
QY 141 CCAAGAA 147
Db 145 AAAAAAA 151

RESULT 8

US-09-018-584A-13/c
Sequence 13, Application US/09018584A
Patent No. 623863

GENERAL INFORMATION:

APPLICANT: Schumm, James W.
APPLICANT: Schumm, Jeffrey W.
TITLE OF INVENTION: MATERIALS AND METHODS FOR
IDENTIFYING AND ANALYZING INTERMEDIATE TANDEM
REPEAT DNA MARKERS
TITLE OF INVENTION: REPEAT DNA MARKERS
NUMBER OF SEQUENCES: 147
CORRESPONDENCE ADDRESS:

ADDRESSES: Promega Corporation
STREET: 2800 Woods Hollow Road
CITY: Madison
STATE: Wisconsin
COUNTRY: U.S.A.
ZIP: 53711-5399

COMPUTER READABLE FORM: 3.5 inch, 1.44 Mb

COMPUTER: IBM compatible PC

OPERATING SYSTEM: Windows 95

SOFTWARE: Word 97 (DOS text format)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/018,584A

FILING DATE: 04-Feb-1998

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Grady J. Frenchick

REGISTRATION NUMBER: 29,018

REFERENCE/DOCKET NUMBER: 16026.9180

TELECOMMUNICATION INFORMATION:

TELEPHONE: (608) 257-3501

TELEFAX: (608) 257-2275

INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:

LENGTH: 372 bp

TYPE: Nucleic Acid

STRANDEDNESS: Double

TOPOLOGY: Circular

MOLECULE TYPE: Genomic DNA

HYPOTHETICAL: no
IMMEDIATE SOURCE:
LIBRARY: plasmid, pGem3zf(+)
CLONE: GL32
POSITION IN GENOME:
CHROMOSOME/SEGMENT: 4 qter
US-09-018-584A-13

Query Match 7.7%; Score 31; DB 3; Length 372;

Best Local Similarity 56.3%; Pred. No. 3.7; Indels 0; Gaps 0;

Matches 58; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 36 GGGAGAGAAATCAAAAGAGCCCAAGAAAGAAACCCAGATGAGAAAGAAACGCGAGGA 95

Db 145 GAG 86

QY 96 AATGACTACATTTGAAGAAAGAACTTCAAGATCAAGATAGAA 138

Db 85 AAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 43

RESULT 9

US-08-520-678A-29/c

Sequence 29, Application US/08520678A

Patent No. 5874565

GENERAL INFORMATION:

APPLICANT: Rice, Charles M.

APPLICANT: Kolykhalov, Alexander A.

TITLE OF INVENTION: NOVEL 3' TERMINAL SEQUENCE OF HEPATITIS

TITLE OF INVENTION: C VIRUS GENOME AND DIAGNOSTIC AND THERAPEUTIC USES THEREOF

NUMBER OF SEQUENCES: 39

CORRESPONDENCE ADDRESS:

ADDRESSEE: Howell & Haferkamp, L.C.

STREET: 7733 Forsyth Blvd., Suite 1400

CITY: St. Louis

STATE: MO

COUNTRY: USA

ZIP: 63105

COMPUTER READABLE FORM: disk

MEDIUM TYPE: Floppy

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/520,678A

FILING DATE:

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: Henderson, Melodie W.

REGISTRATION NUMBER: 37,848

REFERENCE/DOCKET NUMBER: 6029-6836

TELECOMMUNICATION INFORMATION:

TELEPHONE: 314-727-5188

TELEFAX: 314-727-6092

TELEX:

INFORMATION FOR SEQ ID NO: 29:

SEQUENCE CHARACTERISTICS:

LENGTH: 260 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-08-520-678A-29

Query Match 7.6%; Score 30.8; DB 2; Length 260;

Best Local Similarity 53.3%; Pred. No. 3.6;

Matches 65; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 35 TGGAGAGATCAAAAGAGCCCAAGAAAGAAACCCAGATGAGAAAGAAACGCGAGG 94

Db 216 TAGAAGTCAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 157

QY 95 AAATGACTACATTTTGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 154

;; TITLE OF INVENTION: METHODS OF USE THEREFOR
;; FILE REFERENCE: 210121.484C2
;; CURRENT APPLICATION NUMBER: US/09/640,173
;; CURRENT FILING DATE: 2000-08-15
;; NUMBER OF SEQ ID NOS: 196
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 16
;; LENGTH: 396
;; TYPE: DNA
;; ORGANISM: Homo sapien
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (1)-(396)
;; OTHER INFORMATION: n = A,T,C or G
US-09-640-173-16

Query Match 7.6%; Score 30.8; DB 4; Length 396;
Best Local Similarity 51.2%; Pred. No. 4.4;
Matches 65; Conservative 0; Mismatches 62; Indels 0; Gaps 0;
QY 21 AAAAATCAGTTGCTGGGAGAGATCAAAAGAGCCCAAGAAAGGAAAGCCAGATGAGGA 80
DB 131 AAAATTTNNCCCCNAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 72
QY 81 AAGAAACCGCAGGAAATGACTTACATTTTCAAGGAAATCTTCAAGATCAAGTAAAGAAAG 140
DB 71 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 12
QY 141 CCAAGAA 147
DB 11 AAAAAAA 5

RESULT 13
US-09-713-550-16/c
;; Sequence 16, Application US/09/713550
;; Patent No. 6617109
;; GENERAL INFORMATION:
;; APPLICANT: Xu, Jiangchun
;; APPLICANT: Stolk, John A.
;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
;; TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF OVARIAN CANCER
;; FILE REFERENCE: 210121.484C4
;; CURRENT APPLICATION NUMBER: US/09/713,550
;; CURRENT FILING DATE: 2000-11-14
;; NUMBER OF SEQ ID NOS: 205
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 16
;; LENGTH: 396
;; TYPE: DNA
;; ORGANISM: Homo sapien
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (1)-(396)
;; OTHER INFORMATION: n = A,T,C or G
US-09-713-550-16

Query Match 7.6%; Score 30.8; DB 4; Length 396;
Best Local Similarity 51.2%; Pred. No. 4.4;
Matches 65; Conservative 0; Mismatches 62; Indels 0; Gaps 0;
QY 21 AAAAATCAGTTGCTGGGAGAGATCAAAAGAGCCCAAGAAAGGAAAGCCAGATGAGGA 80
DB 131 AAAATTTNNCCCCNAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 72
QY 81 AAGAAACCGCAGGAAATGACTTACATTTTGAAGAAATCTTCAAGATCAAGTAAAGAAAG 140
DB 71 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 12
QY 141 CCAAGAA 147
DB 11 AAAAAAA 5

RESULT 14
US-08-520-678A-22/c
;; Sequence 22, Application US/08520678A
;; Patent No. 5874565
;; GENERAL INFORMATION:
;; APPLICANT: Rice, Charles M.
;; APPLICANT: Kolykhalov, Alexander A.
;; TITLE OF INVENTION: NOVEL 3' TERMINAL SEQUENCE OF HEPATITIS
;; TITLE OF INVENTION: C VIRUS GENOME AND DIAGNOSTIC AND THERAPEUTIC USES THEREOF
;; NUMBER OF SEQUENCES: 39
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Howell & Haferkamp, L.C.
;; STREET: 7733 Forsyth Blvd., Suite 1400
;; CITY: St. Louis
;; STATE: MO
;; COUNTRY: USA
;; ZIP: 63105
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/520,678A
;; FILING DATE:
;; CLASSIFICATION: 536
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Henderson, Melodie W.
;; REGISTRATION NUMBER: 37,848
;; REFERENCE/DOCKET NUMBER: 6029-6836
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 314-727-5188
;; TELEFAX: 314-727-6092
;; TELEX:
;; INFORMATION FOR SEQ ID NO: 22:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 356 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
US-08-520-678A-22

Query Match 7.5%; Score 30.4; DB 2; Length 356;
Best Local Similarity 48.8%; Pred. No. 5.6;
Matches 82; Conservative 0; Mismatches 86; Indels 0; Gaps 0;
QY 48 AAAGAAGCCCAAGAAAGGAAACCCAGATGAGGAAAGAAACCGCAGGAAATGACTACATT 107
DB 251 AAGGAAGGAAAGAAAGGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 192
QY 108 TGAAGAAAGAACTTCAAGATCAAGATCAAGATCAAGATCAAGATCAAGATCAAGATCA 167
DB 191 AAAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 132
QY 168 GGAAGAACGAGATGCGAGTGTCTTCTGAAGAAGTGTCTACACTGTAT 215
DB 131 AAACAGGAATGCGCTTAAGAGCGCGAGTGTCTTACCCCAACCTTCAT 84

RESULT 15
US-08-897-126-22/c
;; Sequence 22, Application US/08897126
;; Patent No. 6297003
;; GENERAL INFORMATION:
;; APPLICANT: Rice, Charles M.
;; APPLICANT: Kolykhalov, Alexander A.
;; TITLE OF INVENTION: NOVEL 3' TERMINAL SEQUENCE OF HEPATITIS
;; TITLE OF INVENTION: C VIRUS GENOME AND DIAGNOSTIC AND THERAPEUTIC USES THEREOF
;; NUMBER OF SEQUENCES: 39
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Howell & Haferkamp, L.C.

STREET: 7733 Forsyth Blvd., Suite 1400
CITY: St. Louis
STATE: MO
COUNTRY: USA
ZIP: 63105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/897,126
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/520,678
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Henderson, Melodie W.
REGISTRATION NUMBER: 37,848
REFERENCE/DOCKET NUMBER: 6029-6836
TELECOMMUNICATION INFORMATION:
TELEPHONE: 314-727-5188
TELEFAX: 314-727-6092
TELEX:
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 356 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-897-126-22

Query Match 7.5%; Score 30.4; DB 3; Length 356;
Best Local Similarity 48.8%; Pred. No. 5.6;
Matches 82; Conservative 0; Mismatches 86; Indels 0; Gaps 0;
QY 48 AAAGAAGCCCAAGAAAGGAAACCCAGATGAGGAAAGAAACCGCAGGAAATGACTACATT 107
Db 251 AAGGAGGAAAGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGG 192
QY 108 TGAAGAAACTTCAAGATCAAGATAAGAAAGCCCAAGAGTTTCATCCACTTCTAATCA 167
Db 191 AAAAAAAAAAGAAA 132
QY 168 GGAACACGAGATGGCAGTGGTTCTCAAGAACTGTGCTACACTGTCAT 215
Db 131 AAACACGGAATGGCCCTAAGAGCGCGAGTGTATTACCCCACTTCAT 84

Search completed: April 23, 2004, 03:09:53
Job time : 65 secs

This Page Blank (uspto)

This Page Blank (uspto)

This Page Blank (uspto)